

Beyond the origin

This issue of *Nature* anticipates next year's bicentenary of Charles Darwin's birth and the 150th anniversary of *On The Origin of Species*. We begin here with a look 50 years into the future.

"Creation is not an event that happened in 4004 bc," the geneticist Theodosius Dobzhansky wrote in 1973. "It is a process that began some 10 billion years ago and is still under way." The realization that the processes of biological creation are at once unspeakably old, and in continuous play around us, is one of the greatest discoveries of history. And yet this discovery — unlike that unceasing and ancient creation itself — can be assigned a well-defined and comparatively recent origin in the mid-nineteenth century.

Ideas on the transmutation of forms and the evolution of life have a long history; so, indeed, do Charles Darwin's personal views on the matter, which have provided historians with grist for many mills. (For *Nature's* Darwin coverage in this issue, see page 295, and online at www.nature.com/darwin.) But the way in which Darwin put together evidence and argument in *On the Origin of Species* marked a definitive break, and an undeniable beginning. The book, 149 years old this week, provided for the first time a way of reconciling life's past and present — a way to explain both the staggering diversity of life and its fundamental unity.

That view of life has been enriched and strengthened in the intervening century and a half, and will continue to be so. But the coming decades could also see Darwin's purview expanded in fundamental ways. The discovery of the universality of the genetic code in the 1960s — the same in elephants and *E. coli*, as the French molecular biologist Jacques Monod famously put it — magnificently bore out Darwin's view that life is united in a common descent. But that need not remain the case.

One distinct possibility is that evidence of life beyond Earth will be found by detecting tell-tale features in the spectra of planets orbiting other stars. Although astronomers are hardly likely to be able to observe variation and evolution of that life in the next 50 years, detection alone could provide insight into the frequency of life's origination. And that, in turn, could help illuminate how life came to be



Darwin200

on Earth — a problem that classical Darwinism is hard put to answer.

An even more likely development is that life will be created *de novo* here on Earth. The first experiments in whole-organism synthetic biology, such as the synthetic mycoplasma being worked on at the J. Craig Venter Institute in Rockville, Maryland, will cleave quite closely to the designs already developed by natural selection. But there are already schemes for going further — for using different genetic codes, for example. Although the synthesis of complex organisms might remain the stuff of fantasy for some time (see page 310), new ways of building self-replicating, one-genome, one-cell organisms seem quite plausible. The development of creatures born from an idea, not an ancestor, will undoubtedly provide new insights into evolution, not least because the proclivities of such creatures to evolve will need to be kept in check.

"Creatures born from an idea, not an ancestor, will undoubtedly provide new insights into evolution."

By the time the 200th birthday of *On the Origin of Species* is celebrated, the life under study by science may well no longer be united by common ancestry in the way that all life is today. In that sense, Darwin's view of the world will have been superseded. But whether that life exists around another star or in a bioreactor, it will still evolve, if given leave to, according to the simple and awe-inspiring algorithms of natural selection.

The essay of Dobzhansky's quoted earlier bears the now-famous title "Nothing in biology makes sense except in the light of evolution". That is so close to being an analytical truth — a necessary implication of what life itself is — that we can be certain it will continue to be true into the future. But that certainty in no way limits the diversity and sheer wonder of what we will find on the voyage that Darwin began. ■

Call to action

European scientists who support neuroscience research on primates should tell their politicians why.

In 1996, Giacomo Rizzolatti and his colleagues at the University of Parma in Italy published a finding that many neuroscientists regard as a landmark. By means of electrodes inserted into the inferior frontal cortex of macaque monkeys, they discovered neurons that responded not only to the monkeys' own actions but also to similar actions the subjects observed in other monkeys. There is strong suggestive evidence of similar 'mirror neurons' in humans. Such neurons, now thought to have a role in understanding

others' behaviour and emotions, have stimulated a great deal of fundamental research, as well as hypotheses relating to several cognitive disorders, including autism.

No grants committee could have foreseen the relevance of these fundamental primate experiments to human pathologies. That is precisely why a new directive proposed by the European Commission earlier this month requires action by anyone who thinks such research to be desirable.

The directive's intent, laudable in principle, is to introduce a new baseline of regulation of the use of animals in research across the 27 member countries of the European Union (EU). Standards of care currently vary greatly across the member states. So the draft directive would enforce on every lab a level of regulation already implemented in the countries most protective of research animals' interests. Such

a reform is certainly needed. And indeed, although there is as yet no schedule for discussion and possible amendment by the European Parliament, it seems likely that a directive in some form will sooner or later be agreed by the parliament and the EU Council of Ministers.

For now, however, the urgent question facing the research community concerns the exact form that the directive will eventually take. As currently worded, it poses an immediate and substantial threat to neuroscience research and to the very benefits to human and animal health that the document says it wishes to support. In particular, it imposes major new restrictions on invasive research using primates. For any primate-research proposal to be approved, it would have to be shown that no alternative species might serve, and that there is relevance to serious diseases or conditions, or to species preservation.

Applied research is mostly carried out in species such as rats or mice for nearly all diseases, including neurological or psychiatric disease. Primates are instead important for more basic research — finding out how the brain works. Most of the labs carrying out invasive research on monkeys in Europe, by recording directly from single neurons, for example, study neurobiological aspects of

consciousness such as attention, or decision-making, or how signals are coded in the brain.

Primates such as macaque monkeys are essential for such work given the similarity of their brains to our own. If we don't understand how brains function in health we will not come close to finding out how to fix brains in disease. It is hard to imagine how brain prosthetics and brain-machine interfaces, which are tantalizingly close to realization, could be properly developed without the support of fundamental primate studies.

The current wording of the directive's text could allow a broad permissive interpretation if an ethics committee decided that very basic research was indeed a prerequisite to finding cures to disease. But in the prevailing climate that seems all too unlikely. It is essential, therefore, that any directive clearly and explicitly permits fundamental research on non-human primates other than great apes. It is also essential that Europe's researchers contact their European Parliament members (MEPs) to tell them as much. They can be found here: <http://tinyurl.com/5k32uf>. MEPs will certainly be hearing from opponents to such research. In politics, the number of voices raised does matter. ■

Stem-cell futures

In the changed political climate, US agencies can provide a new kind of leadership.

Given his campaign promises, it seems likely that US President-elect Barack Obama will move quickly after his inauguration on 20 January to lift the Bush administration's restrictions on federal funding for human embryonic stem (ES) cell research. That will be good news for American scientists, if only because they will no longer have to decide between the human ES cell lines best able to answer their questions and the lines for which they can receive federal funds. But perhaps the biggest advantage is that their most important funding agency, the US\$29-billion National Institutes of Health (NIH), can now start to lead from the front.

The agency has a lot of catching up to do. Since 9 August 2001, when President George W. Bush declared that federal funding could only support work with human ES cell lines already in existence, the NIH has largely had to sit on the sidelines while others stepped in to fill the gap. Several US states launched their own initiatives for funding stem-cell research, complete with peer-review panels and regulatory policies. Meanwhile, in a mirade of organization and diplomacy, the International Stem Cell Forum, a working group chaired by the UK Medical Research Council, coordinated scientists across 11 countries to thoroughly compare and characterize some five dozen ES cell lines, only a few of which could be studied using NIH funds.

Indeed, many of the state, private and international organizations operating during the agency's (relative) absence from the field now have experience and expertise that the NIH lacks — not to mention the capacity and willingness to perform tasks that once might have gone to the NIH by default.

Given this reality, the post-20 January NIH should not expect

to take on the kind of primary leadership role for stem cells that it took for projects such as the Human Genome Initiative. But the agency can take the lead in coordinating and facilitating the many stem-cell programmes that are already under way.

As an example, consider the legal complications that can sometimes bedevil research collaborations trying to work across international borders — or even across state lines. Researchers joke that some collaborations require as many lawyers as scientists. The NIH could do much to simplify matters, particularly in helping states to ensure that human stem-cell lines are derived and used under ethical guidelines, including informed consent.

The NIH could also take responsibility for pushing forward the often tedious work needed to address essential questions. How many ES cells would it take to cause a tumour? How can animals be used to predict the behaviour of transplanted cells? When is a cell pluripotent, and to what extent does the pluripotent state vary?

Before the NIH can pitch in, however, it needs to stand back. The agency should review the portfolio of research it already funds to learn where particular effort is needed to push through translational barriers; it should look for fundamental biological questions where progress might have been slowed by researchers shying away from ES cell research. To get this insight, the NIH must reach out to other governmental and private programmes promoting stem-cell research to discover which strategies have been most successful. Those entities, some of which tirelessly tout their successes in public, must also be willing to reveal their mistakes and limitations. With this knowledge in hand, the NIH can then decide how to foster existing efforts, what new projects to establish, and how it can guide scientists in its own labs to fill in the gaps. ■

"The NIH should not expect to take on the kind of primary leadership role for stem cells that it took for other projects."

RESEARCH HIGHLIGHTS

ZOOLOGY

Whiff of danger

Anim. Behav. doi:10.1016/j.anbehav.2008.08.022 (2008)

The house finch (*Carpodacus mexicanus*) has joined the short list of birds that can smell, apparently using this sense to detect predators.

Timothy Roth and his colleagues at Indiana State University in Terre Haute placed 51 wild finches in endosures in which food sat in ventilation gaps. For two days, unscented air streamed through these gaps and the finches happily moved into the airflow to eat. Roth and his team then gently blew air over cat faeces, rabbit faeces or water before wafting it into the endosures.

Both faecal odours reduced the amount of time the finches spent at feeders. The effect was strongest for cat scent, suggesting that finches can tell a predator from its poo.

PHYSICS

Flags and drag

Phys. Rev. Lett. **101**, 194502 (2008)

Professional cyclists can reduce their drag by following close behind the rider in front. Counterintuitively, downstream flags feel more drag than those flapping in front of them.

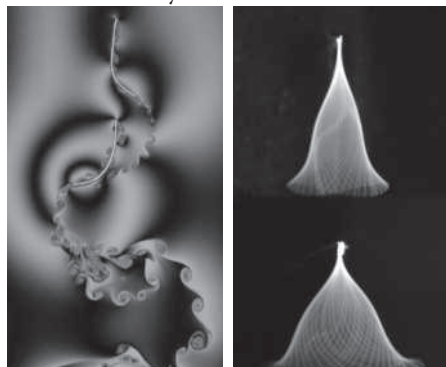
Leif Ristroph of Cornell University in Ithaca, New York, and Jun Zhang of New York University in New York City simulated one-dimensional flags by tethering tiny threads in a flow of soap film (pictured below). The upstream thread's flapping was constrained by the proximity of the fixed 'flagpole' — the leading portion — of its downstream partner. Meanwhile, the downstream flag was caught in the leader's oscillating wake.

CHEMICAL BIOLOGY

Honey trap

J. Am. Chem. Soc. **130**, 15242–15243 (2008)

A nanoparticle with antibody-like binding activity has been created by a group of researchers led by Kenneth Shea of the



AM. PHYS. SOC.



I. SPENCE/FIPA

Quick change

Nature Genet. doi:10.1038/ng.253 (2008)

By mutating just two genes, researchers in Belgium have turned a slender annual plant (inset) into a bushy, woody perennial (main picture).

Perennials have repeatedly

evolved from annuals, but genetic changes that can cause this switch had not been identified. Siegbert Melzer of Ghent University and his colleagues studied *Arabidopsis thaliana* with mutations in *soc1* and *ful*, which both regulate when

the species flowers.

Mutants with dysfunctional forms of both genes formed woody stems and structures resembling dormant buds. These buds awakened periodically, producing several waves of growth and a longer lifespan.

University of California, Irvine. Their plastic antibody' attaches to a specific biological molecule — the honeybee toxin, melittin, which is composed of 26 amino acids.

The authors took a selection of acrylamide monomers containing chemical groups that bind to different portions of melittin, and linked these into a polymer using the target as a template. Crosslinking the monomers generated particles 30–40 nanometres in diameter, making the fruits of their labour comparable in size as well as binding activity to a real antibody. Molecules such as this might one day be injected into the blood, adding new sensors or antidotes to the immune system's armoury.

MOLECULAR BIOLOGY

Silent guide

Cell **135**, 649–661 (2008)

A whole-genome scan of the mouse has revealed where to find the protein H2AZ, which influences gene expression. It's attached to the 'on/off switches' of genes that are

important in embryonic development and tends to accompany the 'polycomb group' protein Suz12, which does a similar job.

Laurie Boyer of the Massachusetts Institute of Technology in Cambridge and her co-workers report that H2AZ can be found at 8% of these switches, or promoter regions, of protein-coding genes in embryonic stem cells. Their gene-silencing experiments showed that H2AZ is needed for embryonic stem cells to become specialized; specifically, the protein must be redistributed around the genome before the cells can acquire a fate.

ASTRONOMY

Missing no more

Astrophys. J. **688**, 277–289 (2008)

Rest assured, astronomers: the missing satellite galaxies of the Milky Way are there — they just haven't been discovered yet, according to calculations by Erik Tollerud of the University of California, Irvine, and his colleagues.

Models that aim to capture how the

clumpiness of dark matter has changed since the Big Bang suggest that the Milky Way should have hundreds, if not thousands, of neighbouring nuggets of dark matter thick enough for faint dwarf galaxies to form in them. But only 20 or so of these galaxies have been spotted.

By assessing the limits of the Sloan Digital Sky Survey, Tollerud and his co-workers have found that the number of known dwarf galaxies is consistent with the number of nearby dark-matter dumps predicted by models.

NANOTECHNOLOGY

Super speakers

Nano. Lett. doi:10.1021/nl802750z (2008)

Imagine a loudspeaker that is bendy, transparent and stretches to twice its length without breaking or changing the intensity of the sound it amplifies. This is what Kaili Jiang and Shoushan Fan of Tsinghua University in Beijing, China, and their colleagues have made. By attaching one of their prototype films to the screen of an iPod they have used it to play sound from the device.

The invention exploits the thermoacoustic effect, which was first realized in the nineteenth century with the platinum thermophone. The effect describes what happens when a material is heated and cooled, thus expanding and contracting the air around it, which creates sound waves. The film on the researchers' iPod generates sound 260 times more efficiently than this historical gadget because the sheets of parallel carbon nanotubes of which it is composed heat up and transfer heat to air faster.

CHEMISTRY

Snip it with zinc

J. Am. Chem. Soc. doi:10.1021/ja801794t (2008)

The covalent bonds that string DNA's nucleotide subunits together are tough, so breaking them requires high temperatures or a catalyst. Now Fabrizio Mancin and Paolo Scrimin at the University of Padova in Italy and their colleagues describe how to make a DNA-cleaving catalyst that works 100 times faster than one of the most efficient synthetic alternatives.

Their method relies on self-assembling sulphur-containing ligands. These are spread across a surface of gold nanoparticles and attach many zinc-based catalysts to each nanoparticle. Because this generates many catalytic sites close together, the substance can cut both strands of DNA at the same point, which enzymes find easy but artificial catalysts have not previously managed.

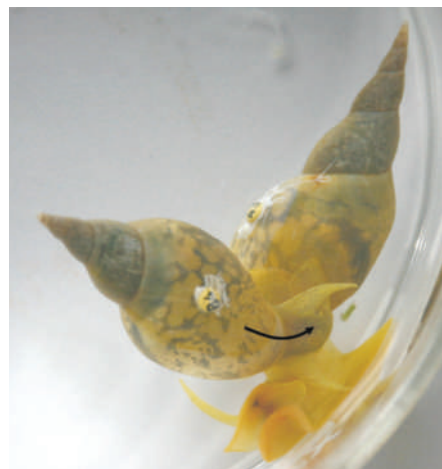
MICROBIOLOGY

The one and only

Science 322, 1110–1112 (2008)

All previously discovered cyanobacteria can use water as an oxygen donor in photosynthesis, producing molecular oxygen as a by-product. But Jonathan Zehr at the University of California, Santa Cruz, and his colleagues have found one that can't.

The new cyanobacterium, known as UCYN-A, has not been cultured, although significant parts of its DNA have been sequenced. It lacks photosystem II, which is necessary for water splitting, but retains photosystem I, in common with light-fuelled but non-oxygen-producing green sulphur bacteria. UCYN-A does not fix carbon dioxide into carbohydrates, but it does fix nitrogen.



A. DAVISON

ZOOLOGY

Sex with a twist

Biol. Lett. doi:10.1098/rsbl.2008/0528 (2008)

The direction in which a snail's shell twists affects much more than the animal's appearance. Angus Davison and his colleagues at the University of Nottingham, UK, have shown that the 'handedness' of great pond snail shells corresponds to the handedness of their courtship.

When hermaphrodites of the species *Lymnaea stagnalis* mate with each other, one takes a 'male role' by climbing onto the other and circling (pictured above) before copulation. All 46 dextral males studied by Davison's team circled anticlockwise during courtship; all but one of 48 sinistral snails circled clockwise.

The researchers say that sinistral snails tend to have mirror-image brains, so the handedness of their behaviour is probably a direct consequence of the handedness of their brains. This probably also dictates the direction of twist in their shells.

JOURNAL CLUB

Jagadeesh Bayry
INSERM, Paris

An immunologist applauds a protein that prunes intolerant white blood cells.

Spreading tolerance is a worthy cause. In the body, newly made white blood cells are rendered tolerant to the many thousands of native proteins. But, like any complex process, this one is not foolproof, and when it goes wrong intolerant white cells cause autoimmune disease.

One way that the tolerance-spreading system can fail is by not having enough 'field agents' to pick off intolerant dissenters. Regulatory T lymphocytes (T_{reg}), a type of white blood cell, are these field agents. They find and suppress other white cells that react to healthy parts of the body. Although it is known that people with low T_{reg} levels tend to have autoimmune diseases, how the cells function has been unclear. Recently, however, researchers in Japan shed light on this mystery.

Kajsa Wing, now at the Karolinska Institute in Stockholm, and her colleagues focused on the protein CTLA-4, which is preferentially expressed by T_{reg} cells and forms part of a rheumatoid arthritis drug called Abatacept. They bred mice without CTLA-4 on the surface of their T_{reg} cells. The animals appeared healthy until maturity, then quickly developed autoimmunity. So CTLA-4 is needed for the field-agent system to operate, and merely expressing it in smaller quantities on other sorts of white blood cell isn't enough. Wing *et al.* then discovered that CTLA-4 on T_{reg} cells interacts with and diminishes two proteins, CD80 and CD86, on the surface of dendritic cells, which show other white cells what to hunt (K. Wing *et al. Science* 322, 271–275; 2008).

All of this confirms that CTLA-4 should provide a means of treating autoimmune diseases. Blocking CTLA-4 should improve the capacity of dendritic cells to present dangerous native cells to the immune system. Clinical trials for cancer treatments that do just that are already under way. Now we have a clearer idea how they work.

Discuss this paper at <http://blogs.nature.com/nature/journalclub>

NEWS

Nuclear renaissance plans hit by financial crisis

Role of fission in fighting climate change looks likely to wane.



Frozen global credit markets and the prospect of a prolonged economic recession are threatening the resurgence of nuclear power that has been touted by governments and industry around the world.

The growing difficulties in attracting investment may prevent nuclear power from playing a significant role in the fight against climate change. A global energy outlook published on 12 November by the International Energy Agency, an intergovernmental organization that guides energy policy, called for an 80% increase in the world's current nuclear capacity by 2030 in order to keep global carbon dioxide levels from rising above 450 parts per million.

That would require bringing some 25 new reactors online every year between now and 2030, five times the current rate of construction, according to Matthew Bunn, a nuclear-policy expert at Harvard University. Given the bleak financial outlook and the limited production capabilities of power-plant vendors, "nuclear can no longer support climate-change needs and targets", he says.

The global downturn means that many utility companies will be "pressing the pause button" on new nuclear plans, says John Gilbertson, who tracks financing for nuclear projects as a managing director for the New York-based brokerage firm Goldman Sachs. "Markets right now are less receptive and capital is more expensive," he says. "Every one of these companies is going through a massive gut-check." Indeed, the French nuclear supplier Areva has told *Nature* that it expects orders for new nuclear power stations to slow, particularly in the United States, where utility companies have applied to the government's Nuclear Regulatory Commission to build seven Areva reactors. But those have yet to translate into confirmed orders with the French company.

Even before the crisis, commercial financing was probably not available in the United States "at any price" without some government incentives, according to Joe Turnage, a senior vice-president for strategy at UniStar Nuclear

Energy, a consortium seeking to build new nuclear plants in Maryland, Pennsylvania and New York state.

But what was a tough task before the downturn is becoming almost impossible in the global recession, according to Mycle Schneider, an independent nuclear consultant based in Paris. Nuclear-parts suppliers are unlikely to expand production, and private utility companies — which must provide billions in investment for new plant construction — have seen their stocks battered in recent weeks. Together with a shortage of skilled labour, the slowdown means that over the next decade "the industry will not even be able to replace the units being shut down because of ageing", predicts Schneider.

Missing targets

The bleak economic outlook makes it even less likely that nuclear will contribute to meeting governments' greenhouse-gas emission targets.

Britain, for example, is about to establish in law a commitment to an 80% reduction in greenhouse-gas emissions, compared with 1990 levels, by 2050 (see *Nature* **456**, 13; 2008). To achieve

those figures would require roughly doubling the current number of nuclear plants in Britain to about 40 by 2025, according to David King, director of the Smith School of Enterprise and the Environment at the University of Oxford, UK, and formerly UK chief scientific adviser.

King believes the target is still achievable, but John Large, an independent, London-based nuclear consultant, disagrees. The shortages in nuclear-plant components created by the downturn will make a radical expansion impossible, he predicts. Consequently, "I suspect we're going to miss our carbon targets".

The outlook, of course, varies from country to country. Some developing nations, notably China, are rolling out major expansions in nuclear capacity. China is planning to more than quadruple its present nuclear output, with a stated goal of 40 gigawatts of installed capacity by 2020. Because this is funded directly by the government, it is less likely to be affected by the current credit crunch.

"Nuclear can no longer support climate-change needs and targets."



But in many other countries where power markets are in various stages of deregulation, utilities must raise billions from investors to build a new plant, even though they are being helped by subsidies and incentives from governments.

With private investors running scared, government financing is becoming ever more vital. For example, the United States Enrichment Corporation (USEC), which enriches uranium fuel for nuclear reactors, announced on 4 November that the credit crunch may force it to delay opening the US\$3.5-billion American Centrifuge Plant in Piketon, Ohio. The plant is one of the few nuclear facilities in the United States that is currently under construction, and was expected to begin operation in 2010.

Although the company was hoping to receive federal loan guarantees, it was also seeking private investment in the project. "The credit market situation simply makes that more

H. HUOVILA/ATVO



HAVE YOUR SAY
Comment on any of our
news stories, online.
www.nature.com/news

Web data predict flu

Two new studies hint at the public-health and research potential of mining the data created as people search the web. Both teams have successfully detected the onset of US seasonal flu epidemics, by extracting patterns of flu-related search terms from the billions of queries stored by Google and Yahoo.

The work tested the hypothesis that people will more frequently search the Internet using flu-related terms when they get sick. One group used Google's search-query logs, the other Yahoo's. Together they generated strikingly concordant findings: patterns of searches matched almost perfectly with official flu surveillance data — and often weeks in advance of these.

The Yahoo research¹, done in conjunction with researchers at the University of Iowa in Iowa City and Harvard University, manually selected keywords for testing such as 'flu' or 'influenza' while eliminating confounding terms such as 'avian' or 'bird'.

The researchers compared the relative frequency of the search terms, between 2004 and 2008, with weekly national data on the standard surveillance indicators of flu — positive viral isolates and flu mortality rates. "We found that we could explain weekly variation in seasonal influenza one to three weeks in advance of cultures, and five weeks in advance of mortality," says Philip Polgreen, an epidemiologist at the University of Iowa.

The Google research, in collaboration with the US Centers for Disease Control and Prevention (CDC), modelled combinations of terms to find best fits against weekly CDC data from between 2003 and 2007. It then created and tested predictive models against the number of influenza-like cases reported by the CDC physician surveillance network during the 2007–08 flu season². The models based on best-fit terms, the researchers claim, were "consistently one to two weeks ahead" of the CDC surveillance reports.

Although search-term analyses could not replace surveillance monitoring — they could be prone to false alarms, and do not provide key information found through recording case numbers and viral isolates — they could supplement existing methods.

In particular, using logs of billions of search queries is fast, cheap and seems to yield good time and spatial resolution, at least down to the state level. Google.org, the company's



Construction of new nuclear power plants is expected to slow as the global recession starts to bite.

difficult," says USEC spokeswoman Elizabeth Stuckle. With private investment so scarce, the company hopes that \$2 billion in loan guarantees from the government will fill the gap.

UniStar also hopes to secure guarantees from the US government, so that it can take out affordable loans from the Federal Financing Bank, a government corporation that oversees borrowing from the US Treasury. Together with incentives from France and Japan, which will supply parts for the new reactors, Turnage believes his consortium can begin construction of a 1,600-megawatt reactor in Maryland in mid-2009. Even in the current turmoil, "our best judgement now is that they're a compelling investment", he says.

How compelling such an investment appears will depend enormously on the policies of President-elect Barack Obama, who has

advocated a cap-and-trade system to regulate carbon dioxide emissions, although without a specific commitment to nuclear energy to help reach those goals.

Schneider says he is still sceptical about whether a new wave of nuclear power can really get off the ground. "There are loads of plans, but the situation has worsened considerably," he says. "It's not going to happen."

However, King notes that major public-works projects are one way for governments to prime their economies. The economic downturn shouldn't necessarily slow down nuclear expansion, he says, "provided that the heads of state get it right".

Geoff Brumfiel

For more coverage of what the global financial crisis means for science, visit www.nature.com/financialcrisis

philanthropic arm, has turned its findings into Google Flu Trends, a free daily flu forecast for the United States. It plans to roll out the service to other countries, languages and diseases.

"Google Flu Trends nicely illustrates both the broad coverage and fine-grained detail that one can get from search-engine query-log data," says Jon Kleinberg, a specialist in network analysis at Cornell University in Ithaca, New York.

Scientists have only begun to scratch the surface of the potential of large web data sets — often because companies are reluctant to share their data because of both real and perceived privacy concerns, says David Lazer of the John F. Kennedy School of Government at Harvard University. "It is difficult to get access to such data, and so it should be, because there are confidentiality issues," he says. In the flu-prediction case, this was much less of an issue, says Lazer, as

the data used were large pooled aggregates of search-term frequencies, from which it would be impossible to obtain information on individuals' searches.

In 2006, lacking access to Google databases, Gunther Eysenbach of the University of Toronto took out a Google Ad that pointed users searching for flu-related terms to a flu educational website. The patterns of clicks on the ad correctly

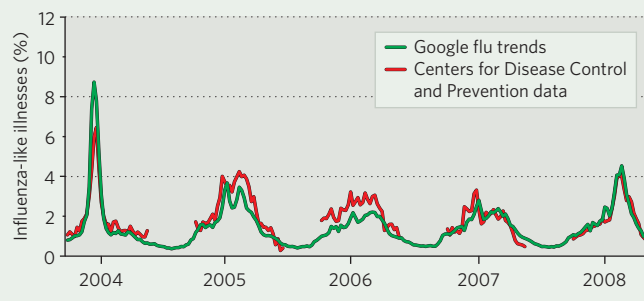
phenomena," he says. That might even include the credit crisis, he says — because banks' own activities give them better and earlier data on macroeconomic trends than those that government agencies can collect.

Declan Butler

1. Polgreen, P. M., Chen, Y., Pennock, D. M. & Nelson, F. D. *Clin. Infect. Dis.* **47**, 1443–1448 (2008).
2. Ginsberg, J. et al. *Nature* doi:10.1038/nature07634 (2008).
3. Eysenbach, G. *AMIA Annu. Symp. Proc.* 244–248 (2006).

TRACKING THE FLU

The relative frequency of flu-related keywords in Google searches closely tracks flu statistics in the US mid-Atlantic region as monitored by government officials.



Gene testing of embryos needs guiding

Couples might soon use the technologies sold by personal-genomics companies to choose the genetic make-up of their children.

Last week, at a meeting of the American Society for Reproductive Medicine in San Francisco, California, researchers cautioned that they do not yet advocate this use of genetic tests. But as the technology advances, consumer demand is likely to overwhelm societal ethical qualms. "If people think they can make their babies healthier at year one, year two, or *in utero*, they will do it," says Jacques Cohen, research director at Reprogenetics, a genetic-testing company in West Orange, New Jersey.

Fertility doctors already use genetic analyses to check cells from embryos for specific large genetic abnormalities, such as the loss or gain of entire chromosomes. These analyses help parents choose which embryos to transfer into a woman's uterus during *in vitro* fertilization procedures. But now scientists have begun more detailed genetic

analyses that rely on microarrays to survey an embryo's genome. Some of the microarrays contain thousands of tiny genetic markers called single nucleotide polymorphisms, or SNPs, that are used to reconstruct the embryo's chromosomal make-up.

The 'SNP chips' used in these tests, which are sold in the United States and could soon be available in the United Kingdom, are similar to those used by personal-genomics companies. That means they could potentially be used to estimate an embryo's risk of adult-onset diseases — such as diabetes and obesity — and other traits, such as hair, eye and skin colour.

Professional societies and governments have not yet decided how to handle preimplantation testing for the risks of such diseases, because it has never before been possible. The UK Human Fertilisation and Embryology Authority, for example, has allowed clinics to use preimplantation genetic diagnosis to screen for serious adult-onset

diseases caused by single genes, such as cancer syndromes, and for genes that are likely but not certain to cause disease, such as the breast cancer risk gene *BRCA1*.

But most common adult-onset diseases, such as diabetes, have multiple genetic contributors — most of which are still unknown. And SNPs only give a very rough estimate of a person's risk of a disease; they do not determine that an embryo will grow into a person who will actually get that disease.

Reproductive geneticists seem wary of treading on such controversial ground. "We have this information," says William Kearns, director of the Shady Grove Center for Preimplantation Genetics in Rockville, Maryland. "But just because we have it doesn't mean we should use it." Kearns's centre will begin offering SNP-based analyses of chromosomal abnormalities next month, but will not report risk information about multi-gene complex diseases to prospective parents.

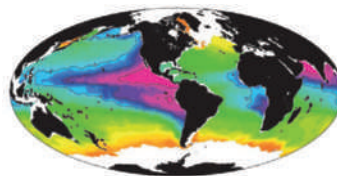
Neither will Gene Security

Network, a company in Redwood City, California, that has already begun selling microarray-based preimplantation genetic diagnosis tests of chromosomal abnormalities and some diseases to fertility clinics. "Any disease we screen for is nasty, early-onset and ethically uncomplicated," says Dave Johnson, Gene Security Network's director of research and development.

Alan Handyside, the inventor of preimplantation genetic diagnosis and director of the London Bridge Fertility, Gynaecology and Genetics Centre, UK, is also developing a microarray-based test, which he hopes will enter a clinical trial in the next several months.

Cohen says that as the understanding of disease genetics progresses, use of tests that seem controversial now may become more acceptable in the future: "If you had the chance to decrease your child's risk of a disease like diabetes and you didn't, society would blame you."

Erika Check Hayden



CLIMATE CHANGE
Marine dead zones set to
expand rapidly.
www.nature.com/news

IFM-GEOMAR

Forestry carbon dioxide projects to close down

The US Department of Energy has decided to close several large-scale, open-air projects designed to evaluate how increased concentrations of carbon dioxide affect forests. The move is necessary, many researchers say, to free up funds for a new generation of more sophisticated experiments.

Free-Air CO₂ Enrichment (FACE) experiments pump carbon dioxide into the air to raise the local concentration of the gas in forests, grasslands and agricultural fields. Forest sites are particularly ambitious and expensive: Ram Oren, lead investigator of Duke University's FACE project in Durham, North Carolina, estimates that every year the experiment's four giant rings of carbon-dioxide-emitting pipes use bottled gas costing US\$900,000. The site, which began operating in 1996 and is the oldest of the forest FACE projects scattered around the world, has been used to study tree growth and productivity under conditions of at least 550 parts per million of carbon dioxide — nearly 1.5 times higher than present levels.

But lead researchers at two FACE sites — Duke's loblolly pine experiment and the Oak Ridge National Laboratory's sweetgum plantation in Tennessee — were told in September to prepare for closure around 2011. Other sites are already closing (see table for examples).

On one level, the closures are painful for a scientific community that craves long-term data. "I'm puzzled," says Oren. "You build FACE experiments to study ecosystem processes, and some of those are very slow." But many also view them as necessary given limited research budgets.

The Department of Energy spends up to \$3 million a year on the Duke site, \$2 million of which goes to Brookhaven National Laboratory in New York for supplies of carbon dioxide and engineering support. In 2006, an agency



The closure of Duke University's FACE project could be used to fund a new generation of experiments.

advisory committee recommended the closures, citing infrastructure limitations. At Oak Ridge, the trees have grown taller than the towers that support the gas pipes; at Duke, trees in the rings have outgrown the surrounding forest, creating an 'edge effect' that could affect experimental results.

In April, the energy department sponsored a workshop to discuss the design of the next generation of FACE experiments. Those attending were hopeful that future sites could vary not only carbon dioxide concentration but other factors as well, including temperature, precipitation and nutrients. Researchers are also thinking about testing different ecosystems, particularly tropical and high-latitude forests, which are thought to be among the biggest players in the terrestrial carbon cycle.

Each of these goals will face technical hurdles. Cheap sources of carbon dioxide may be hard to find in remote areas. This could pose a particular problem in the tropics, where the bigger trees will require larger gas supplies, notes Franco Miglietta of the National Research Council's Institute of Biometeorology in Florence, Italy. And the technology does not yet exist to increase forest temperatures uniformly from root to canopy.

Meanwhile, the Duke and Oak Ridge FACE experiments continue to yield unexpected results. Richard Norby, a lead investigator on the Oak Ridge project, was nearly ready to end the project a few years ago. At that stage, he says, the data seemed consistent from year to year: higher carbon dioxide levels led to increased plant productivity. But then he noticed that the trees' response to the gas had begun to decline. Norby thinks that nitrogen limitation might be the culprit, and Oak Ridge teamed up with Duke in 2005 to study this effect in both forests.

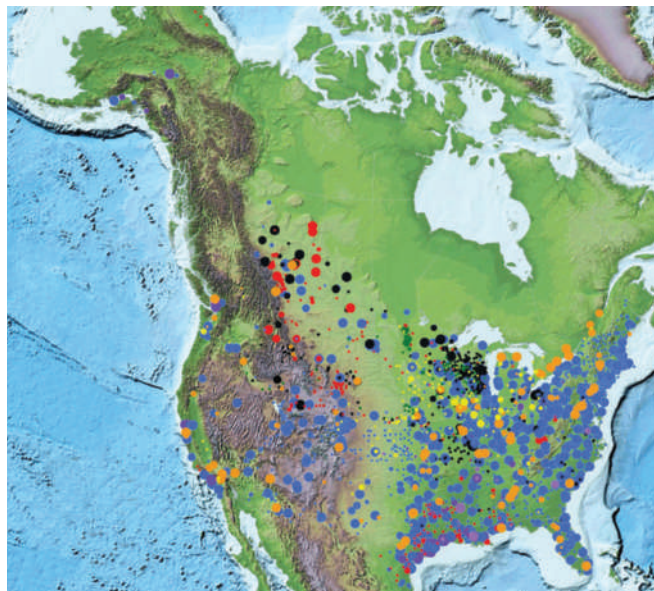
Those experiments will not be completed in time for the planned shutdown dates. But that is part of the difficulty with setting the finish line for a long-term project, says Michael Kupperberg, programme manager for the projects at the energy department's Office of Science. "I've been around scientists my whole professional career," he says, "and I've never had one come to me and say 'I'm done'. There are always new experiments around the corner."

Heidi Ledford

OPEN-AIR CARBON DIOXIDE EXPERIMENTS, PAST AND PRESENT

Site	Ecosystem	Plot diameter	Started	Current status
Durham, North Carolina	Loblolly pine (<i>Pinus taeda</i>) forest	30 metres	1996	Closure in 2011-12
Rhinelander, Wisconsin	Aspen (<i>Populus</i> spp.) forest	30 metres	1997	Closing by 2011
Oak Ridge, Tennessee	Sweetgum (<i>Liquidambar styraciflua</i>) plantation	25 metres	1997	Closing by 2011
Mojave Desert, Nevada	Desert scrub (<i>Larrea tridentata</i> , <i>Ambrosia dumosa</i> , <i>Lycium</i> spp.)	23 metres	1997	Harvesting now
Viterbo, Italy	Poplar (<i>Populus</i> spp.)	20 metres	1999	Closed in 2005

C. HILDRETH/DUKE UNIVERSITY PHOTOGRAPHY



SNAPSHOT Carbon stores

A newly updated atlas from the US Department of Energy shows how North America might store its excess carbon dioxide — if it could figure out how to do so.

On the left is a map showing many of the known stationary

CO₂ sources in the United States and Canada. Blue represents electricity generation, with many power plants clustered in the east and midwest. Orange represents cement manufacturing, including groupings in southern California. Red indicates petroleum and natural-gas processing, showing up most starkly as the Alberta tar sands in Canada.



Potential reservoirs for sequestering much of the CO₂ from such activities could be deep saline formations (above right, in blue), or layers of rock permeated with brine. The new atlas estimates that saline formations could hold anywhere from 920 billion–3,400 billion tonnes of CO₂. This dwarfs the potential for existing oil and gas

reservoirs (82 billion tonnes) or unmineable coal seams (156 billion–184 billion tonnes).

Yet turning saline formations from dream reservoir into sequestration reality remains a challenge. In Alberta, a consortium is in the process of identifying potential formations for a test project to begin in 2009.

Alexandra Witze

Electron ‘bump’ may confirm dark matter

A high-altitude balloon experiment above the Antarctic seems to have seen a possible signature of mysterious ‘dark matter’, similar to that spotted earlier this year by a European satellite.

The Advanced Thin Ionization Calorimeter (ATIC), an experiment to search for charged particles from space, has spotted a surplus of high-energy electrons coming from somewhere in the cosmos (see Letter, page 362, and News & Views, page 329). Although the interpretation is far from certain, the electrons could be produced by dark matter — previously undetected particles that physicists believe make up 85% of all matter in the Universe.

ATIC’s findings are similar to data from the PAMELA (Payload for Antimatter Matter Exploration and Light-nuclei Astrophysics) satellite mission, a collaboration between Italy, Russia, Germany and Sweden that spotted an excess of high-energy positrons, or anti-electrons, at

roughly similar energies (see *Nature* 454, 808; 2008). “In several respects the two measurements complement each other,” says John Wefel, head of the ATIC collaboration and a physicist at Louisiana State University in Baton Rouge.

Wefel’s team looked at data from two multi-day ATIC missions flown between 2000 and 2003 some 35 kilometres above the Antarctic



Several experiments may have found dark matter, the existence of which is inferred here (blue).

ice. With 99.5% of Earth’s atmosphere beneath them, the missions measured electrons that come from various galactic sources such as exploding stars. As predicted by theory, the experiment saw fewer electrons at higher energies. But between 300 gigaelectronvolts and 600 gigaelectronvolts, the number of electrons rose sharply before falling off to background levels.

PAMELA’s results, as posted on the arXiv preprint server last month (O. Adriani *et al.* <http://arxiv.org/abs/0810.4995>; 2008) and also submitted to *Nature*, show positrons increasing up to 100 gigaelectronvolts. Although the collaboration did not report beyond this energy level, some suspect that the number of positrons may continue to increase at higher energies. And, because the PAMELA and ATIC data were measured in different ways, researchers believe that their findings probably confirm each other.

**METHANE ON MARS**

Engineers rule out gassy landing spot as too risky.

www.nature.com/news

NASA/JPL-CALTECH/UNIV. ARIZONA

The 'bump' ATIC sees in the number of electrons could be the result of heavy dark-matter particles annihilating, according to Wefel. When this happens, two dark-matter particles collide and their mass is converted to pairs of fast-moving electrons and positrons, the energies of which would correspond to the mass of the original particles.

"That's certainly the sexiest of the possibilities," says Dan Hooper, a theoretical physicist at Fermi National Accelerator Laboratory in Batavia, Illinois. The exact nature of the dark-matter particles that produce electrons is uncertain, but one idea is that they may be ordinary particles that spend part of their lives in a compact extra dimension of space. Whereas the particles would appear relatively stationary to observers trapped in three spatial dimensions, they could be moving at ultra-high speeds in a fourth spatial dimension. At high speeds, they would create a gravitational force that could be felt by matter trapped in three dimensions of space-time. "It's very wild," Hooper says.

Other, more mundane reasons might also explain the data. The leading candidate is a nearby pulsar, says Aldo Morselli, a particle physicist from the Italian National Institute of Nuclear Physics at the University of

Rome, Tor Vergata. Pulsars, the fast-spinning remnants of supernovae explosions, have enormous magnetic fields that can accelerate electrons to the high energies seen in the experiments, although ATIC's bump is harder to match to a pulsar than PAMELA's increase.

Wefel agrees that it is too early to say for certain whether ATIC and PAMELA have seen dark matter. "You cannot make a hard and fast claim," he says. "The case is still open."

But more evidence may come soon. NASA's Fermi telescope (formerly known as GLAST), which launched in June this year, is designed to hunt for high-energy γ -rays. But the telescope can also spot electrons and positrons, according to Morselli. In the coming months, he predicts that the telescope will be able to verify the positron and electron data of PAMELA and ATIC. In addition, Fermi may be able to spot γ -rays that have come from dark-matter annihilations. "By the spring of 2009, we'll have a lot more information," Morselli says. ■

Geoff Brumfiel

Middle East synchrotron on the lookout for funds

The Middle East's first synchrotron facility, a collaborative effort intended to promote peace through international scientific cooperation, formally opened its doors on 3 November in Allan, Jordan. But unless the project can plug a €15-million (US\$19-million) shortfall, the machine's completion risks being stalled.

Following a flurry of behind-the-scenes discussions, officials are confident that they will find the money. The facility's inauguration at Al-Balqa Applied University may have helped: when ministers and other representatives finally saw the buildings, and the beamlines and other machinery being installed, it had "a big psychological impact," says Herwig Schopper, who retired as head of the project council on 4 November.

The idea of a synchrotron as a bridge for peace in the Middle East was first proposed in 1997; two years later, SESAME (Synchrotron-light for Experimental Science and Applications in the Middle East) launched as an autonomous international research organization, under the auspices of the United Nations Educational, Scientific and Cultural Organization (UNESCO). It is modelled on CERN, Europe's particle-accelerator centre near Geneva, Switzerland. Jordan won the bid to host it.

Most of SESAME's technical equipment comes from donations from facilities being replaced or upgraded elsewhere. Its 0.8 gigaelectronvolt (GeV) injector system comes from BESSY I in Berlin, Germany, and entire beamlines from the Daresbury Synchrotron Radiation Source in the UK, the Laboratoire pour l'Utilisation du Rayonnement Électromagnétique (LURE) in Orsay, France, and the Swiss Light Source in Villigen. In addition, powerful magnets and other equipment are being loaned by the Lawrence Berkeley National Laboratory and the SLAC National Accelerator Laboratory, both in California. SESAME is supposed to be completed by 2011, with a price tag of roughly \$100 million, which includes labour and the value of donated parts, says Chris Llewellyn-Smith, the new head of the SESAME council.

But instead of the 1 GeV machine initially proposed for the site, SESAME members decided

several years ago to opt for an ambitious 2.5 GeV machine, comparable to new third-generation sources elsewhere. The project members — Bahrain, Cyprus, Egypt, Israel, Iran, Jordan, Pakistan, the Palestinian Authority and Turkey — agreed to pay for the increase in operational costs, which will rise progressively from \$1.3 million to \$5 million annually over the next four years. The problem is funding the capital costs of the 2.5 GeV storage ring.

Several possible solutions seem to be emerging. Jordan is said to be offering a loan to prevent delays. As the host country, it has



Open SESAME? The facility in Jordan awaits further funding.

already contributed more than \$25 million in land, buildings and overheads, and cannot donate more directly without effectively making SESAME a national project.

Sources close to the discussions say that Iran is contemplating offering to build those parts of the ring for which it has the technical capacity. This would represent a big reduction in the €15-million gap. Qatar, a small but wealthy country, seems likely to join the project, as does Iraq, and possibly other countries in the region. But the wealthiest country in the area, Saudi Arabia, has so far declined to join, because it does not collaborate with countries with which it does not have diplomatic relations — and that includes SESAME member Israel.

Llewellyn-Smith says he will also approach philanthropies and wealthy individuals. "It's obvious to anyone who goes to SESAME now that this is going to happen," he says. "We need to ensure it's finished quickly, and that the scientific project is first-rate."

Declan Butler

Satellite risks losing sight of Earth

NASA is scrambling to work out what to do with a US\$100-million Earth-science satellite that has been sitting on a sanitized shelf in Greenbelt, Maryland, for nearly eight years.

Last month, President George W. Bush signed a NASA reauthorization bill that, among other things, ordered the agency to come up with a plan for the Deep Space Climate Observatory (DSCOVR), which should have been launched around 2001. The space agency is now in talks with the National Oceanic and Atmospheric Administration (NOAA) and the US Air Force about finally getting the probe off the ground. But the negotiations might mean that the spacecraft loses its Earth-observing instruments and instead goes into orbit with a remit to stare only at the Sun.

Stripping the two Earth-monitoring systems from DSCOVR to save money is an “appalling” idea, says Francisco Valero, the mission’s principal investigator at the Scripps Institution of Oceanography in La Jolla, California. A 2006 NOAA study estimated it would take \$117 million to refurbish and launch the spacecraft; a 2007 study from NASA put that figure at \$205 million.

DSCOVR was meant to be the first Earth-observing mission sent to the Lagrange-1 (L-1) point, a parking place 1.5 million kilometres away from Earth in the direction of the Sun. From there the craft would take measurements that would allow scientists to work out Earth’s energy budget — how much radiation the planet absorbs and re-emits — in one take.

Sun studies

Satellites in low-Earth orbit make similar energy measurements but can observe only small sections of Earth at a time. DSCOVR would offer a “global, rather than myopic, perspective of the planet”, Valero says. One of its Earth-monitoring instruments, a spectroradiometer, would indirectly measure variables such as ozone levels, aerosols, cloud thickness and water vapour. The other, a radiometer, would measure reflected and emitted radiation for the whole planet.

Yet the Air Force and NOAA seem more interested in studying the Sun, Valero says. Both are responsible for monitoring solar weather and storms, which can damage or knock out communications and global-positioning satellites. The Advanced Composition Explorer (ACE) satellite is currently based at



NASA

An Earth-observing satellite that can see the whole planet is described as ‘crucial’ to climate research.

L-1 to watch for solar eruptions. It sends a one-hour advance warning back to Earth of any impending storms.

The 11-year-old ACE instruments are nearing the end of their designed lifespan, so “we are reviewing a variety of options” to replace them, says Colonel Shawn Barnes of the Air Force Space Command. DSCOVR has three Sun-observing instruments that could perform the ACE observations, and top NASA officials have confirmed that the three agencies are discussing how to use the mothballed craft to monitor solar wind and space weather.

“DSCOVR would offer a global, rather than myopic, perspective of the planet.”

DSCOVR was originally meant to fly around 2001 aboard the space shuttle. But NASA postponed the launch after Congress demanded an independent review of the probe’s scientific merit. The

satellite is a renamed version of Al Gore’s proposed Earth-watching satellite, called Triana. After the review found that the mission’s science goals were still valid, NASA sought another ride for it. In 2001, the French space agency discussed with NASA the possibility of launching Triana, but funding constraints ruled that out. And the 2003 Columbia disaster meant that the mission was out of luck for a shuttle ride. NASA eventually cancelled

the mission in January 2006.

In May 2007, NASA held a workshop to look at whether NOAA might fly the spacecraft instead. NOAA wanted the launch but “DSCOVR didn’t rise to be NASA’s highest priority”, says Hal Maring, the NASA programme manager who directed the workshop. That’s in part because the 2007 National Research Council decadal survey did not rank the mission among the 17 ‘highest priority’ Earth-observing missions that NASA and NOAA should develop and launch in the next ten years.

Since then the scientific community has showed renewed interest in seeing DSCOVR fly. In March 2008, the Ernst Strüngmann Foundation in Frankfurt, Germany, held an invitation-only forum for 44 top climate scientists. Many participants, none of whom was directly involved with DSCOVR, agreed that satellite observations of Earth from L-1 are essential for assessing changes in cloud cover and climate.

The issue was pushed to the top of the agenda last month after Bush signed the NASA Reauthorization Act, legislation that directs the space agency in its programmes and budget priorities for 2009. NASA has until 13 April next year to report its plans to Congress. ■

Ashley Yeager

Coal-fired power plants face delay in United States

In a case that could have wide-ranging implications for future US power plants, a federal appeals board has ordered the Environmental Protection Agency (EPA) to reconsider a decision not to regulate carbon dioxide emissions from a coal-fired plant in Utah.

On 13 November, the EPA's Environmental Appeals Board rejected environmentalists' claims that the agency must regulate CO₂ emissions from a 110-megawatt power plant planned by the Deseret Power Electric Cooperative. But the board also rejected the EPA's argument that it cannot do so because of historical precedent.

The case hinges on a 2007 Supreme Court ruling that the EPA has the authority to regulate greenhouse-gas emissions from automobiles. Environmental groups contend that the authority should extend to the power sector, too.

Recognizing the precedent, the appeals board recommended that the EPA address the question at the national level. President-elect Barack Obama's campaign has said it would pursue such authority if Congress does not act on global warming.

US court allows sonar exercises to continue

The US Supreme Court has ruled that national security trumps environmental law, at least in the case of whales and sonar. Following the decision, the Navy no longer has to stop sonar training exercises off the southern California coast if whales, dolphins and other marine mammals are spotted nearby.

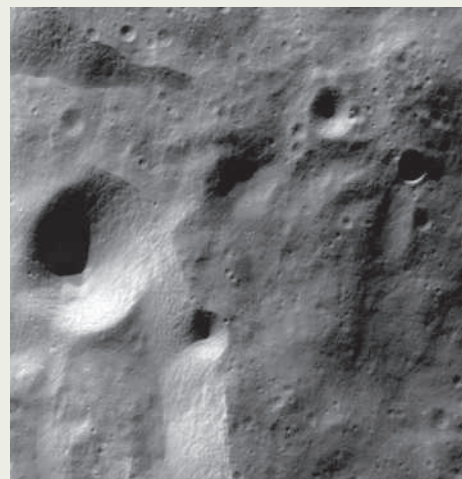
In a 5–4 ruling released on 12 November, the court rejected arguments that the restrictions were necessary to protect the animals from becoming harmed or disoriented by the sonar. The concerns are “plainly outweighed by the Navy's need to conduct realistic training exercises to ensure that it is able to neutralize the threat posed



Whales will have to put up with the US navy's training exercises.

Indian space mission stakes its claim on the Moon

A small probe launched from the Indian Space Research Organisation's (ISRO) craft Chandrayaan-1, now in orbit around the Moon, slammed into the Shackleton crater close to the Moon's south pole on 14 November. The probe, about 40 centimetres wide, sported the Indian flag painted on its sides and collected data during the descent that ISRO says will help to plan a future landing mission. Chandrayaan-1, India's first unmanned lunar mission, will circle 100 kilometres above the Moon's surface for the next two years. This image was taken by its Terrain Mapping Camera from lunar orbit the day after releasing its impactor probe, and shows the bright rim of the 117-kilometre-wide Moretus crater at the Moon's south pole.



ISRO

by enemy submarines,” Chief Justice John Roberts wrote for the majority opinion in *Winter v. Natural Resources Defense Council*.

For a longer version of this story, see <http://tinyurl.com/6jvxor>.

Drug giant shows its commitment to stem cells

On 14 November, pharmaceutical company Pfizer announced the formation of its new stem-cell research unit, Pfizer Regenerative Medicine, which will be based in two locations on either side of the Atlantic. The site in Cambridge, UK, will focus on neural disorders, whereas the base in Cambridge, Massachusetts, will explore applications of stem-cell technology to endocrine and cardiac research.

The centres will employ about 70 scientists in total, and Pfizer plans to announce additional collaborations with academia and industry in coming weeks.

Pfizer already has a long-term interest in cell-based therapeutics, but its expanded commitment to stem-cell research is indicative of a recent investment trend by other large pharmaceutical companies eager to use stem cells for drug screening, including GlaxoSmithKline, based near London, and the Swiss companies Roche and Novartis.

Ecstasy could augment the benefits of psychotherapy

The controversial drug MDMA — also known as ecstasy — can help ease the effects of post-traumatic stress disorder, according to the first phase-II clinical trial into the potential therapeutic benefits of using the drug as an adjunct to psychotherapy.

Most patients in the trial who were given psychotherapy along with doses of MDMA (3,4-methylenedioxy-N-methyl-

amphetamine) experienced statistically significant reductions in the severity of their condition after two months, compared with a control group who received psychotherapy and a placebo. The results, which have yet to be submitted to a peer-reviewed scientific journal, were presented on 13 November at the International Society for Traumatic Stress Studies convention in Chicago, Illinois.

“It's important to realize this is a small pilot study and it will be necessary to replicate the results elsewhere,” says Michael Mithoefer, a private-practice psychiatrist based in Mount Pleasant, South Carolina, who led the study, funded by the Multidisciplinary Association for Psychedelic Studies, an advocacy group based in California. “But it is evidence that this should be studied further,” he adds.

For a longer version of this story, see <http://tinyurl.com/6b6zhp>.

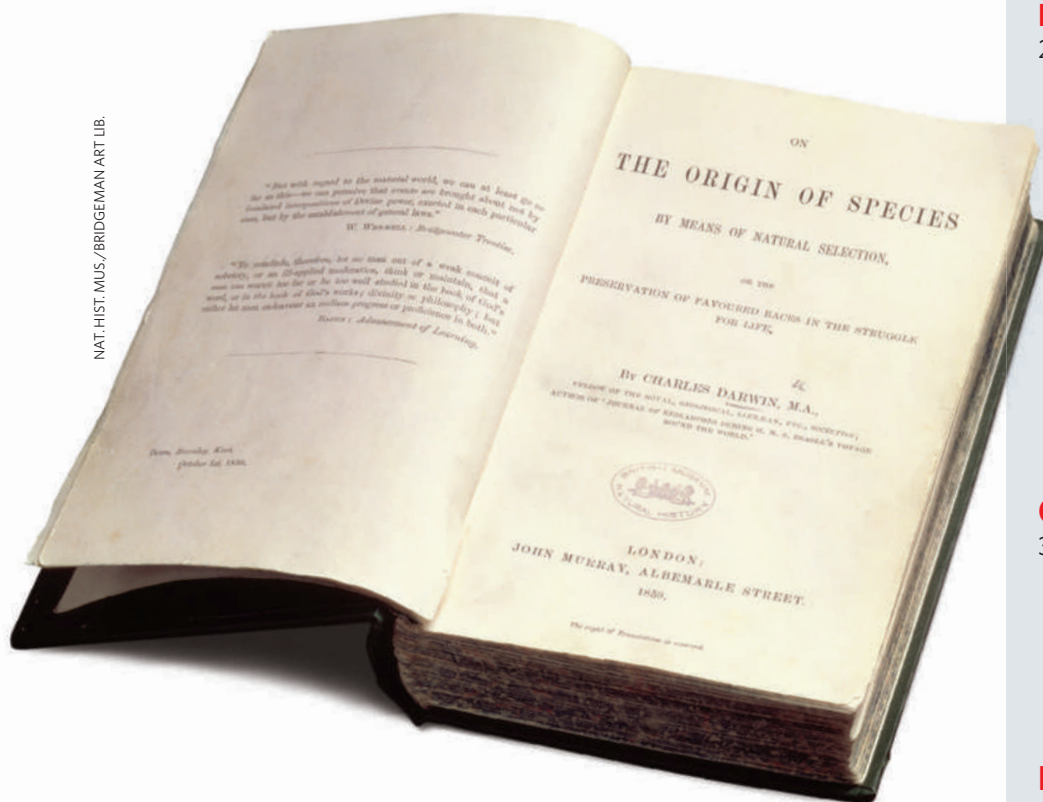
US drug agency opens outposts in China

In a bid to regulate imported food and medical products from China, the US Food and Drug Administration is opening three inspection offices in Beijing, Shanghai and Guangzhou this week — its first such offices outside the United States.

The agency will also set up inspection offices in India, Europe and Latin America by the end of the year, with another office to follow in the Middle East in 2009. The move signals the country's increasing attempts to tighten the regulation of imported products.

Several health scares over Chinese-made products have triggered mounting safety concerns in the United States. In the latest incident, melamine and related compounds — which contaminated Chinese dairy products, killing at least four children and causing tens of thousands more to fall ill — were found in food imports from China.

NAT. HIST. MUS./BRIDGEMAN ART LIB.



BEYOND THE ORIGIN

Celebrating the man and the book



Darwin200

Charles Darwin did not merely open a new chapter in the story of biology; he opened a new book. The publication of *On the Origin of Species* in 1859 was an event without any parallel, the sudden and powerful exposition of a new view of life as mutable, competitive and shaped by its environment laid out in the form of one long argument.

As we enter the year of Darwin's 200th anniversary, this week's *Nature* celebrates the impact of *On the Origin of Species* by looking back and looking forward. Janet Browne offers a guide to how the 50th and 100th anniversaries of its publication were marked, and Marek Kohn looks at the long-running debate over how, and to what extent, selection can work on groups rather than individuals. *Nature* readers tell us of their ambitions for the coming year of Darwin celebrations, and we provide a guide to when and where some of those celebrations are taking place.

We also look forward to the promise of new research into how natural selection shapes not just the forms creatures take but also the remarkably diverse ways they reach those forms. And we ask what stands in the way of re-originating species by looking at the technology and know-how that would be needed to turn the genome of the woolly mammoth into a living thing.

The ways that evolutionary biology has developed, and the range of areas into which it offers insight, are now almost as diverse as the unending carnival of natural forms that Darwin sought to explain. But they can all be traced back to a single place: the origin. ■

EDITORIAL

281 **Beyond the origin**

GALLERY FEATURE

304 **An eye for the eye**
Simon Ings

NEWS FEATURES

296 **The needs of the many**
Marek Kohn

300 **Beneath the surface**
Tanguy Chouard

310 **Let's make a mammoth**
Henry Nicholls

COMMENTARY

317 **Great expectations**
Patricia Adair Gowaty, Ismail Serageldin, Per-Edvin Persson, Niles Eldredge, Michael Lynch, Masatoshi Nei, Ulrich Kutschera, Mustafa Akyol, Randolph Nesse, Mel Greaves

BOOKS & ARTS

320 **Colonies that conquer**
Manfred Milinski

322 **Darwin: heading to a town near you**
Joanne Baker

323 **A Down House bookshelf**

ESSAYS

324 **Birthdays to remember**
Janet Browne

NEWS AND VIEWS

330 **Mammoth genomics**
Michael Hofreiter

LETTERS

387 **Sequencing the nuclear genome of the extinct woolly mammoth**
Webb Miller, Stephan Schuster and colleagues

395 **Mechanism of phototaxis in marine zooplankton**
Gáspár Jékely and colleagues

PODCAST

Listen to: Simon Ings and Gáspár Jékely on the eye; Marek Kohn on group selection; and Henry Nicholls and Stephan Schuster on the making of mammoths.

For more online see
www.nature.com/darwin





THE NEEDS OF THE MANY

The idea that natural selection acts on groups, as well as individuals, is a source of unending debate. **Marek Kohn** reports on what the two sides disagree about — and why it matters to them.

If biologists have learnt one thing about evolution over the past 40 years, it is that natural selection does not work for the good of the group. The defining insight of modern Darwinism is that selection ‘sees’ individuals and acts on them through the genes they embody. To imagine otherwise, generations of students have been warned, is to fall into a naive error definitively exposed as such in the mid-1960s.

Yet group selection — the idea that evolution can choose between groups, not just the individuals that make them up — has a higher profile today than at any time since its apparent banishment from mainstream evolutionary theory. And it gets better press, too. This is in part owing to the efforts of David Sloan Wilson, of Binghamton University in New York, who argues that the dismissal of group selection was a major historical error that needs to be rectified. And it does not hurt that he has been joined by Edward O. Wilson, the great naturalist and authority on social insects. They and many others have worked to reposition group selection within the broader theme of selection that acts simultaneously at multiple levels.

The remarkable feature of the debate is not that neither side has managed a conclusive victory; it is that they seem, from the outside, to agree on so much. “Everyone agrees that group



Darwin200

selection occurs,” says Andy Gardner of the University of Edinburgh, UK. Yet Gardner and his colleagues Stuart West and Ashleigh Griffin have trenchantly criticized¹ David Sloan Wilson’s arguments on this subject — a critique to which David Sloan Wilson responded by initiating a lengthy debate in the community under the heading ‘If the theorists cannot agree...’.

There is widespread agreement that group selection and kin selection — the post-1960s orthodoxy that identifies shared interests with shared genes — are formally equivalent. “If you’re talking about altruism, the mathematics that gives you kin selection is also the mathematics that if you slice it another way gives you group selection for altruism,” says Andrew Bourke, who studies social insects at the University of East Anglia, UK. To some researchers, this makes the choice between the two frameworks a matter of perspective. Others doubt that the idea of group selection adds much, if anything, to the understanding obtained from kin selection theory. And for theorists such as Gardner and his colleagues, debates about group selection are largely semantic.

But others see deeper conceptual disagreements, which, according to Samir Okasha, a professor of philosophy at the University of Bristol, UK, “stem from underlying theoretical assumptions about how we should ‘carve up’ nature”.

Okasha, who has written a well-regarded book on the subject², reels off a string of questions. How should a ‘group’ be defined? How should ‘altruism’ and ‘selfishness’ be defined? Does the ‘selfish-gene’ view of evolution conflict with the idea that units of selection can act at different hierarchical levels? Does group selection require that groups engage in a form of reproduction, as individuals do, or just that groups must have differing levels of productivity?

To clarify that last issue, Okasha distinguishes between two kinds of multi-level selection, known in the trade as MLS1 and MLS2. MLS1 focuses on how the division of a population into groups affects the frequencies of different types of individuals in the population. The classic example would be altruism: a population divided into groups may reward altruism more than an undifferentiated one would. MLS2 focuses on the frequencies of the groups themselves. The fittest groups in the MLS1 sense are those that contribute the most individuals to the next generation; the fittest groups in the MLS2 sense are those that contribute the most groups.

It is MLS1 that is implicated in the question at the heart of debates over levels of selection — how altruism can spread through a population. And it is MLS1 that can be taken as equivalent to kin selection. Interest in MLS2 processes has grown more recently, as the question of what levels of selection there are has become

A. NUP SHAH/NATUREPL.COM; M. SPENCER GREEN/AP



entwined with that of how those levels evolved; how the advent of multicellularity, say, created an 'organism' level above the level of the cell. In such major evolutionary transitions, smaller units are integrated into larger ones, be it replicating molecules into simple cells or organisms into societies. For a stable new class of entity to emerge, selfish tendencies among their constituent units had to be suppressed at the group level.

Establishing priorities

Charles Darwin saw that what was good for the group might not be good for the individual. In *The Descent of Man* (John Murray, 1871), he observed that "although a high standard of morality gives but a slight or no advantage to each individual man and his children over the other men of the same tribe", it would "certainly give an immense advantage to one tribe over another". However, by the 1960s, many naturalists and biologists had come to see the relationship between what is good for the individual and what is good for the group as unproblematic. They assumed that individuals would subordinate their own interests to those of the species.

Some, though, had doubts. And when the late naturalist Vero Wynne-Edwards argued explicitly that natural selection acted on groups and that individuals restrained their consumption and reproduction for the good of the community³, his book provoked a small but powerful backlash. An under-appreciated alternative model for the same phenomena had recently been proposed by the theorist Bill Hamilton, then at the London School of Economics. Hamilton argued that the spread of altruistic behaviour, which Wynne-Edwards assumed would be selected because of its benefits to the group at large, would in fact be determined by the degree to which altruists shared genes with

their beneficiaries. A quartet of the critics met in Oxford and coined the term 'kin selection' to distinguish Hamilton's theory from Wynne-Edwards's group selection.

Kin selection was easily and powerfully expressed in mathematics, and became the new orthodoxy, its ascendancy cemented in place by its forceful and compelling popularization in Richard Dawkins's *The Selfish Gene* (Oxford Univ. Press, 1976). Group selection might be possible in theory, these thinkers allowed, but it could be ignored in practice. As the theorist George C. Williams declared in his 1966 book *Adaptation and Natural Selection* (Princeton Univ. Press) — a rallying call against group selection that had great influence in America — "the higher levels of selection are impotent".

With the focus on altruism, disadvantageous by definition, the impotence of the higher levels seemed insurmountable. Evolutionary models developed by John Maynard Smith, the most junior member of the Oxford quartet, drove home the point that it would take just one selfish individual to spoil things for everybody. Cheats would inevitably multiply at the altruists' expense. Selection for characteristics that benefited a group at the expense of individuals would be opposed by selection within the group, and selection at that lower level would win.

In their recent writing on the subject⁴, Wilson and Wilson accept the first part of this proposition, but challenge the second. Sometimes selection between groups is weak and sometimes it is very strong, they argue; the balance between levels of selection should be evaluated case by case. Drawing on Darwin's original insight, they boiled down the principles underlying the altruism question to three short sentences. "Selfish-

ness beats altruism within groups. Altruistic groups beat selfish groups. Everything else is commentary."

Gardner, Griffin and West distinguish between the 'old' group selection of Wynne-Edwards and this 'new' version, embedded within a framework of multi-level selection⁴. David Sloan Wilson, however, just sees "one long argument" since the 1960s. Group selection was originally rejected on the grounds that it was implausible rather than on evidence it didn't happen, he says. But the evidence has changed. "It can now be accepted on the best of evidence," says Wilson, "not that it always happens, but it happens some of the time."

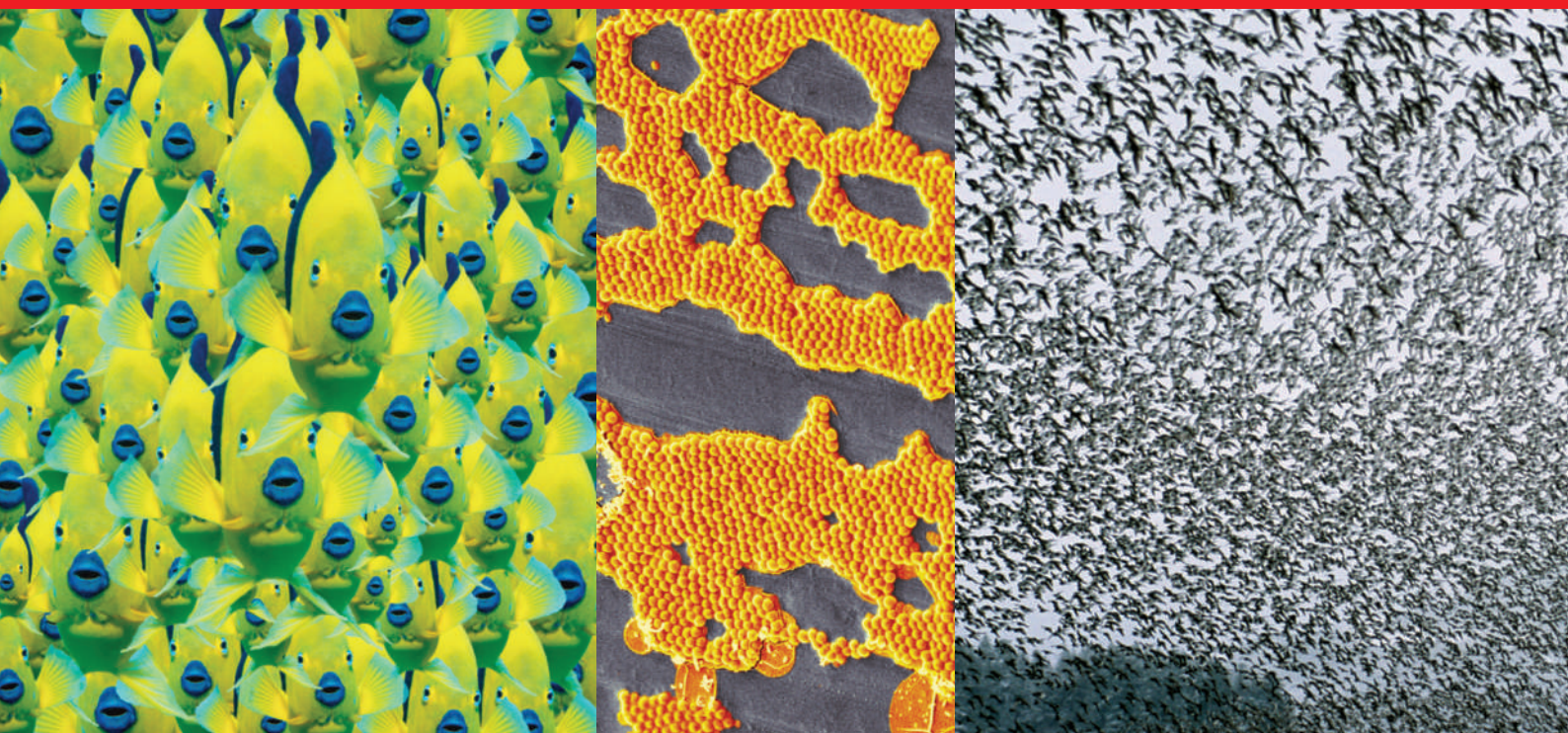
Wilson highlights experiments on microbes as being "totally definitive" in this regard — even if the experimenters don't always present their results that way. Benjamin Kerr, of the University of Washington in Seattle, and his colleagues have shown, for example, that "prudent"

"The adaptation of individual organisms is an amazing result. It should be treasured rather than abused."
— Andy Gardner

viruses, which do not over-exploit their bacterial hosts, prevail over "rapacious" ones when their ability to disperse and find new prey is restricted⁵. Wilson and Wilson say that the study "elegantly establishes the plausibility of Wynne-Edwards' hypothesis ... without citing Wynne-Edwards or the term group selection".

Kerr is happy to affirm the connection. Wynne-Edwards observed that groups would evolve to conserve their resources for the benefit of their descendants if those descendants lived in the same place — or as he put it², "fall heirs to the same ground". Kerr recalls that in 1964 Maynard Smith discussed the prospects for group selection on this basis using a model that involved imaginary mouse colonies in haystacks⁶. "Our experiment was similar in spirit to Maynard

T. DAVIS/CORBIS; DIGITAL VISION/GETTY IMAGES; D. GULIN/GETTY IMAGES



R. A. CLEVELAND, CORBIS; SCIMAT/SPL; M. QUINTON/GETTY IMAGES

Smith's intent," says Kerr: "to elucidate the conditions favouring the type of restraint Wynne-Edwards wrote about." The intent is indeed similar in that the model and the experiment are attempts to put Wynne-Edwards's verbal arguments into a more rigorous form. But whereas Kerr feels that Wynne-Edwards had a point, Maynard Smith actually wanted to dispose of an argument he regarded as silly. The haystack model established the view that conditions permitting group selection were highly restrictive and unlikely to be encountered in real life.

David Sloan Wilson also points to a study⁷ by Paul Rainey and Katrina Rainey that looked at bacteria cooperating to form a mat on the surface of a liquid medium. The mat is made by mutants that produce larger than normal quantities of a sticky polymer, a trait called 'wrinkly spreader' (WS). Wilson and Rainey say that selection between groups maintains the WS trait, even though the individuals making the polymer incur costs that the other bacteria don't have to pay. But Rainey, who works at the New Zealand Institute of Advanced Study at Massey University in Auckland, says that this is not what the experiment showed. "In fact, given our experimental design, this possibility is more or less ruled out. Kin selection provides the simplest explanation for the evolutionary emergence of the WS type." After all, the fact that daughter cells are physically stuck to their progenitors tends to keep things in the family.

Paul Rainey does not dispute that kin selection is equivalent to multi-level selection of the MLS1 type. Indeed, he has described the emergence of the mats as "a triumph of selection acting at the higher level".⁸ He is quick to point out, however, that this is not group selection of the MLS2 type.

In Rainey's eyes, David Sloan Wilson does not provide an adequate solution to the problem that MLS2 selection needs groups to reproduce

as individuals. Wilson proposes that selection may act on groups identified by shared traits rather than kinship, and that such groups may be small and ephemeral. Rainey considers such models useful for explaining how cooperation is maintained, but doubts that they can account for adaptation at the level of the group, and is especially sceptical that they can explain the evolution of group-level adaptations during a major evolutionary transition such as the integration of cells into multicellular organisms. That requires an explanation for how selection shifts its focus from individuals within groups to groups themselves; from MLS1 to MLS2. Theorists have not paid much attention to this transition, Rainey says. "Samir Okasha refers to this as a grey area. I think it is a great black box!"

Rainey thinks that his mat work could help to open up that black box, and that the individuality groups need for MLS2 selection might develop in such systems. Cells that 'cheat' by enjoying the benefits of being in mats without contributing the chemical that holds them together may in fact have a role at a higher level. Because they can swim away and give rise to new colonies, they have the potential to act as a kind of germ line for colonies that otherwise cannot reproduce. If cheats can work like this they could give the mats the individuality that is needed if natural selection is going to operate at the colony level.

There is a problem here. For a cheat to found a colony, a fresh mutation is needed — WS has to be rediscovered. Rainey, though, is confident that selection can impose design on the process. From the outset, selection will work to optimize the rates at which cheaters arise — groups with too many fall apart. Rainey predicts that with this toe-hold in the process, evolution will

produce systems that can switch between mat forming and swimming stages by changing gene expression, not waiting for *de novo* mutation.

Rainey affirms that as natural selection favours the evolution of cheats whenever cooperation has a cost, selection will act on any population that contains cooperative types at both the individual and the group level. But he cautions that this should not make group selection the default explanation for the evolution of the higher level. His reservation echoes a complaint made by Bourke. "I don't like what I see as an almost systematic downplaying of the evidence in favour of kin selection," Bourke says. "There are lots of other people out there,

working on vertebrates right through to microbes, who find evidence for kin selection."

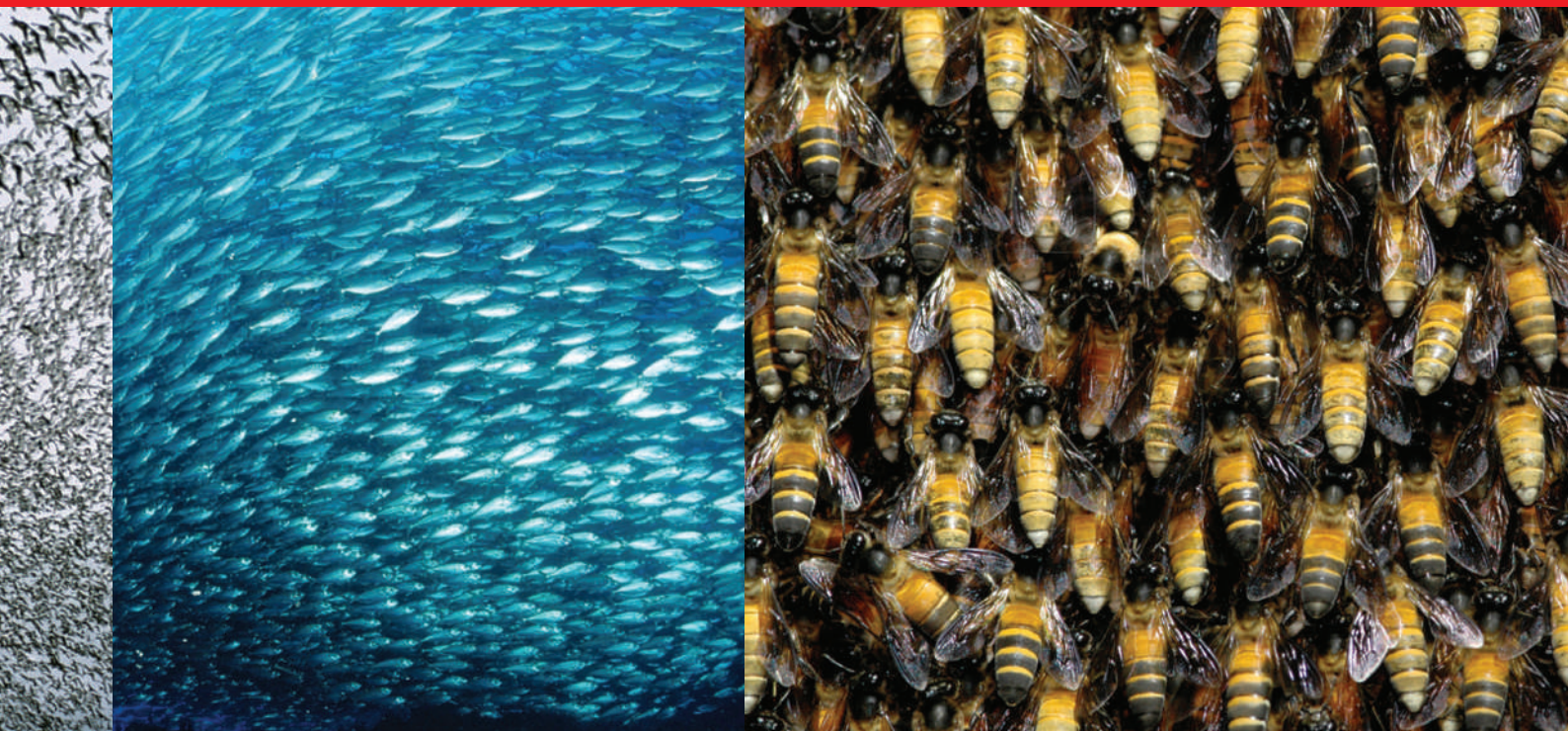
Bourke states the orthodox case firmly. "Everything is gene selection," he declares. When populations are structured into groups, either kin- or group-selection ideas may

need to be applied. "But it is still gene selection. What matters is the kind of trait being selected for, and the structure of the population it is in." If a group-level selection effect is seen, it will be because of relatedness — a nonrandom association of genes. "I don't see that labelling it group selection adds anything to that insight, which was essentially Hamilton's." Kerr, by contrast, regards kin selection and group selection not as alternative views but as potentially complementary perspectives. "Looking at the same system in different ways can often provide deeper understanding," he says.

Although semantic and conceptual confusion are in ample supply, on their own they don't sustain the controversies over group selection. Core beliefs are at stake, and powerful intuitions are at work.

"Our species is the primate equivalent of a beehive or a single organism."

— David Sloan Wilson



From an orthodox point of view, group-selection ideas get in the way of an appreciation of the power of natural selection. Gardner says that in principle he is equally happy with kin selection and multi-level selection, but in practice he prefers to work with the former. It is closer to the basic theory of adaptation, connected through Hamilton to Ronald Fisher, one of the architects of the 'modern synthesis' that brought together population biology and evolutionary theory in the middle of last century⁹. It keeps the focus on the constantly striving individuals that populate Darwin's vision.

And the idea of adaptation is to be treated with respect, not carelessly applied at the level of the group, where it doesn't really work. "I guess the reason I fell in love with biology rather than, say, physics, is that you have this purely mechanical process — natural selection — that gives rise to purpose and function and the appearance of design in the living world," Gardner says. "The adaptation of individual organisms is an amazing result, and nothing like it occurs elsewhere in the natural sciences. It should be treasured rather than abused."

Avoiding the implications

Group-selection thinking is perceived by some as not just an abuse of natural selection but also a denial of it. Tellingly, Dawkins chose evolutionary theorist Robert Trivers's version of this argument for his recent anthology¹⁰ of great science writing. Imagining that evolution works for the greater good looks like a way to avoid the radical implications of natural selection — like claiming that it has little effect, for instance, on the human mind.

For David Sloan Wilson, however, it is the individualistic perspective that betrays influences from outside science. "The concept that self-interest can be a grand explanatory principle seemed to become current in the middle

of the twentieth century, and what happened in evolutionary biology was part of that trend," he says. "Now, against the background of individualism and self-interest, we have the proposition that our species is the primate equivalent of a beehive or a single organism — so what are the implications of that? That's not science fiction; that is the growing consensus."

And this human dimension is ultimately why group selection arouses passions. When Darwin considered selection at group levels, he thought of social insects and of humans. When Hamilton extended the equation derived by his collaborator George Price to accommodate group-selection processes, he combined the discussion with speculations about the adaptive value of warfare as a way for groups to compete — and encountered a reception similar to that of a cat who presents its owner with a half-dead mouse. But that was more than 30 years ago. Today the idea that warfare and altruism are two sides of the same coin — that competition between groups by way of warfare encourages altruism within groups — is widely discussed, as promulgated¹¹ by Samuel Bowles of the Santa Fe Institute in New Mexico.

Robert Boyd of the University of California, Los Angeles, and Peter Richerson at the University of California, Davis, have promoted the idea that selection between cultural groups can be highly effective in humans. Whereas genetic-group models are vulnerable to the fact that individuals can migrate from group to group, reducing the between-group variance on which selection works, cultural models have the advantage that human migrants can adopt the cultures of the groups to which they migrate. People intuitively recognize their species as 'groupish' and realize that models of group evolution can have implications for the understanding of human societies.

That is part of their attraction for Joel Peck,

whose modelling studies at the University of Sussex, UK, suggest that the conditions under which group selection could work might not be as restrictive as his late colleague Maynard Smith had implied. Underlying his interest in the life sciences, Peck says, is a desire to be involved in positive social change. "Group selection is one of the topics within evolutionary biology that is most directly about how societies change." He notes that in the major transition that unified the ancestors of the mitochondrial and nuclear genome into the eukaryotic cell, the two ancestors' interests were almost entirely united. As well as being fascinating in themselves, processes such as these "may give us hints about how to unify the interests of different human individuals, or groups of individuals", he says.

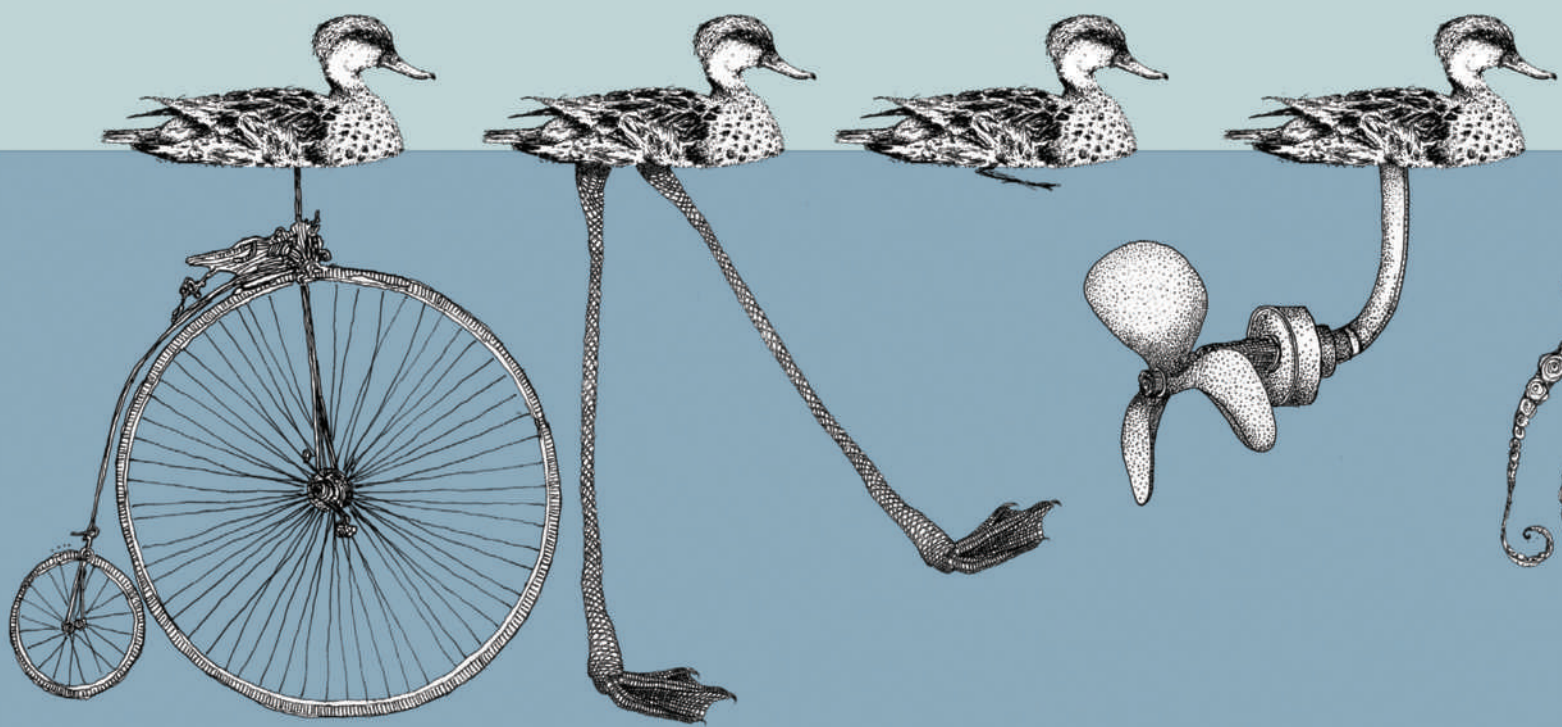
Wilson suggests that in the light of group selection, the human species displays unity; Peck hopes that group selection may show humans how to achieve unity. Ultimately, their science is about the good of the species. ■

Marek Kohn is a writer in Brighton, England. He is the author of *A Reason For Everything: Natural Selection and the English Imagination*.

1. West, S. A., Griffin, A. S. & Gardner, A. *J. Evol. Biol.* **21**, 374–385 (2008).
2. Okasha, S. *Evolution and the Levels of Selection* (Oxford Univ. Press, 2006).
3. Wynne-Edwards, V. C. *Animal Dispersion in Relation to Social Behaviour* (Oliver & Boyd, 1962).
4. Wilson, D. S. & Wilson, E. O. *Q. Rev. Biol.* **82**, 327–348 (2007).
5. Kerr, B., Neuhauser, C., Bohannan, B. J. M. & Dean, A. M. *Nature* **442**, 75–78 (2006).
6. Maynard Smith, J. *Nature* **201**, 1145–1147 (1964).
7. Rainey, P. B. & Rainey, K. *Nature* **425**, 72–74 (2003).
8. Rainey, P. B. *Nature* **446**, 616 (2007).
9. Whitfield, J. *Nature* **455**, 281–284 (2008).
10. Dawkins, R. *Oxford Book of Modern Science Writing* (Oxford Univ. Press, 2008).
11. Choi, J. K. & Bowles, S. *Science* **318**, 636–640 (2007).

See also pages 281 and 295, and online at www.nature.com/darwin.

R. DIRSCHERL/FLPA; S. WILLIAMS/NATUREPL.COM



Beneath the surface

You might think that once evolution has found one way to get something done, it will stick with it. But similar physical forms can hide radically different wiring, finds **Tanguy Chouard**.

It is not birth, marriage or death, but gastrulation, which is truly the most important time in your life,” British embryologist Lewis Wolpert famously said. That’s when the primordial sheets of embryonic cells are instructed to fold inwards on their way to becoming more specialized tissues. The process is governed in part by a group of cells called the Spemann organizer, and developmental biologists, echoing Wolpert’s view, thought that you couldn’t mess with something as important as the organizer.

So when, in the early 2000s, those biologists tried to find the Spemann organizer in tunicates they were in for a big surprise. Tunicates — also known as sea squirts — are humans’ closest invertebrate cousins. They have tadpole-like larvae that closely resemble miniature vertebrate embryos and so were expected to build their bodies in the same way. But they don’t. Most of the ‘organizer genes’ are there in the tunicate genome, but they are expressed elsewhere in the embryo and do dramatically different things¹. It’s as if you had found a car in which components of the engine were scattered all over the back seat — but the car still worked.

Many biologists, consciously or not, tend to see living systems as optimally tuned. If one species has a complex solution to a difficult problem (such as a Spemann organizer for



Darwin200

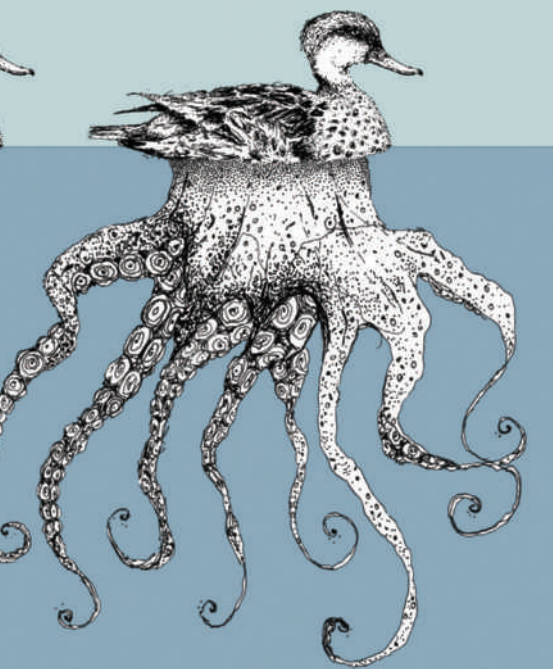
building a swimming tadpole), it is often assumed that the way genes are orchestrating that process will be the same in related species. That idea has only been strengthened by the discovery that many genes, and the proteins they encode, display a stunning degree of conservation across hundreds of millions of years of evolutionary time.

What is becoming clear, however, is that although most genes might be conserved, the regulatory connections that control their expression might not be. Closely related species can connect up their genes in very different regulatory networks, while keeping the end result deceptively unchanged. “Problems with many solutions are the rule rather than the exception in living systems,” says Andreas Wagner, a bioinformatics expert at the University of Zurich, Switzerland. Now researchers are trying to understand how evolution finds the solutions it does, and why. Some think that this ‘underground’ variation was selected for. Some think that it appeared by chance. And regardless of how it arose in the first place, some believe that a variety of regulatory networks may offer an evolutionary advantage in the future. Trying out many different designs ‘underground’ could provide a hidden source of evolutionary innovation, variation on which organisms might draw when faced with new challenges.

All these ideas present a major challenge to those who study ‘evo-devo’, the evolution of developmental processes. Researchers can no longer conclude that two organisms are built in the same way by considering one gene or even one signalling pathway at a time. They must consider the entire system with its inputs, outputs and the connections in between. Where evo-devo has met systems biology, a new discipline — systems evo-devo — has emerged. “I’ve never been too fond of any of those buzzwords,” says Patrick Lemaire at the Developmental Biology Institute of Marseilles in France, “but if this means bringing more rigour to evo-devo, making it less descriptive and more mechanistic, then, yes, a systems framework seems useful.”

Conservative ideology

It is hard to shake the belief that conservation runs deep. From the 1980s onwards, DNA sequencing and gene-knockout technology revealed the extent to which genes and their functions were conserved. These techniques showed that many of the genes that determine the animal body plan are virtually identical in both structure and function in creatures that, on the outside, have little in common. The expression patterns of *Hox* genes, for example, specify the same positional values in the head-to-toe body axis of fruitflies and mice. Such discoveries strengthened the intuitive idea that when



K. LEMON

gene sequence and organization are conserved, so too is gene function — and that conserved function implies conserved genetics. But in the past few years it has become apparent that the gene–function relationship is not so clear cut. A given function can result from diverse combinations of the same genes, or different genes, even in very closely related species.

Molecular dissections of the extremely conserved body plan of insects — head, thorax and abdomen — provide a striking example of how different genes and wiring can lead to very similar endpoints. In the fruitfly *Drosophila*, the *Hox*-related gene *Bicoid* is essential to establishing this form. Work starting in the 1980s has shown that messenger RNA coding for the Bicoid protein is glued to the front end of the egg. Here, the protein is translated after fertilization, diffuses from one end of the embryo to another and orchestrates the activity of several other key genes involved in distinguishing head from tail. In fact, Bicoid does so many wonderful things that it was widely assumed to be fundamental to insect development.

Not so. Starting in the late 1990s, sequence data made clear that the *Bicoid* gene, although crucial in *Drosophila*, is absent from most other

insects. This apparent paradox is resolved only when considering gene regulatory networks as a whole. The labs of Reinhard Schroeder, of the University of Tübingen, Germany, and Claude Desplan, of New York University, for example, found that several genes that are activated by Bicoid after fertilization in fruitflies are deposited into the egg before fertilization in the flour beetle² and a parasitic wasp³. The lack of *Bicoid* is compensated by minor changes in the action of several other genes in the developmental network.

So whether you think about the Spemann organizer or the *Bicoid* gene, the same rule seems to apply: there are many combinations of contributing factors that can reach the same outcome. “You can’t understand any of this if you think at the single-gene level,” says Lemaire. The information that determines biological function lies at a higher, more abstract level, in the entire network of genes, proteins and other factors that each act on the others in a series of nonlinear feedback loops. The body plan or feature that results is what scientists who study complex systems call an ‘emergent property’ — one that is more than the sum of its parts.

It all feels very counter-intuitive. Maybe it is because man-made machines are sensitive to individual component failure, that biologists tend to assume that genetic networks are similarly constrained, and that changing their wiring piecemeal would mess up the whole system. So one of systems evo-devo’s key questions is how the genes and their interactions can change while keeping the output of the system as good as, or better than, it was before. To find out, researchers need an experimental system simple enough that many species could be analysed in the finest molecular details. Sexual differentiation in yeast offers such a system, one conserved across species that diverged up to about a billion years ago.

Most yeasts have two mating types: ‘a’ cells express ‘a’ genes, and ‘α’ cells express ‘α’ genes. In the human pathogen *Candida albicans*,

a-specific genes must be actively turned on by a DNA-binding protein called a2 before they are expressed. But in the brewer’s and baker’s yeast *Saccharomyces cerevisiae*, a-specific genes have to be turned off by a repressor protein α2 — their default setting is on. The circuitry works in completely different ways and *S. cerevisiae* doesn’t even have a copy of the a2 gene vital to *C. albicans*. But these two systems didn’t just fall from the sky, they must have evolved from a common ancestral system.

To find out how, Alexander Johnson and his team at the University of California, San Francisco, teased out the way that a-genes are regulated in 16 yeast species whose genomes had been sequenced and could thus be reliably ordered on a phylogenetic tree⁴. They identified the regulatory DNA controlling many of the a-genes and compared the binding sequences for a2, α2 and a third DNA-binding protein called MCM1, which associates with either a2 or α2 depending on the yeast species. They also looked for changes in the way the proteins interact with each other. They reasoned that changes

in the proteins and their binding sites might reveal how one regulatory system incrementally changes to give rise to the other.

And they did. *Kluyveromyces lactis* — a yeast used in the

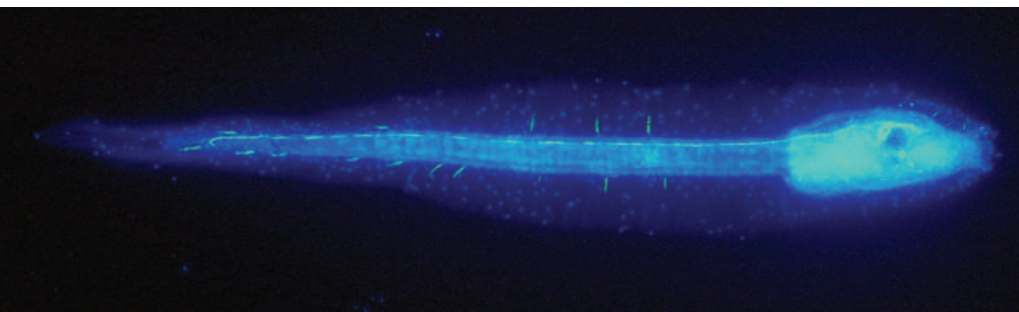
cheese industry — carries what is likely to have been the transition state between the ancestral *C. albicans* logic and the *S. cerevisiae* scheme. Some of its a-genes bind a2, some bind α2 and some bind both. At the same time, MCM1 is able to bind both a2 and α2. “This is absolutely terrific work,” says Sean Carroll, a developmental biologist at the University of Wisconsin in Madison. “They took enough snapshots to show us the whole movie of the evolutionary process.” The key to this overhaul is functional redundancy: by possessing more than one regulatory pathway to achieve the same function, it is possible for one to change while the other keeps the system running smoothly. It’s like moving around a boat in rough weather: you always keep at least one of two clips hooked to the lifeline.

Johnson’s yeast studies revealed how an organism might switch from one regulatory network to another without losing fitness in between. But such experimental studies can only deal with a couple of genes at once. So Wagner and two French collaborators turned to computer simulation to ask how far successive small steps can take a large gene-regulatory network away from its starting point⁵.

Wagner’s model consists of a number of genes, each of which codes for a regulatory protein that activates or represses one or more of the other genes. In the virtual embryo with

“You can’t understand any of this if you think at the single-gene level.”

— Patrick Lemaire



Tunicates such as *Ciona intestinalis* build a tadpole differently from all their vertebrate cousins.

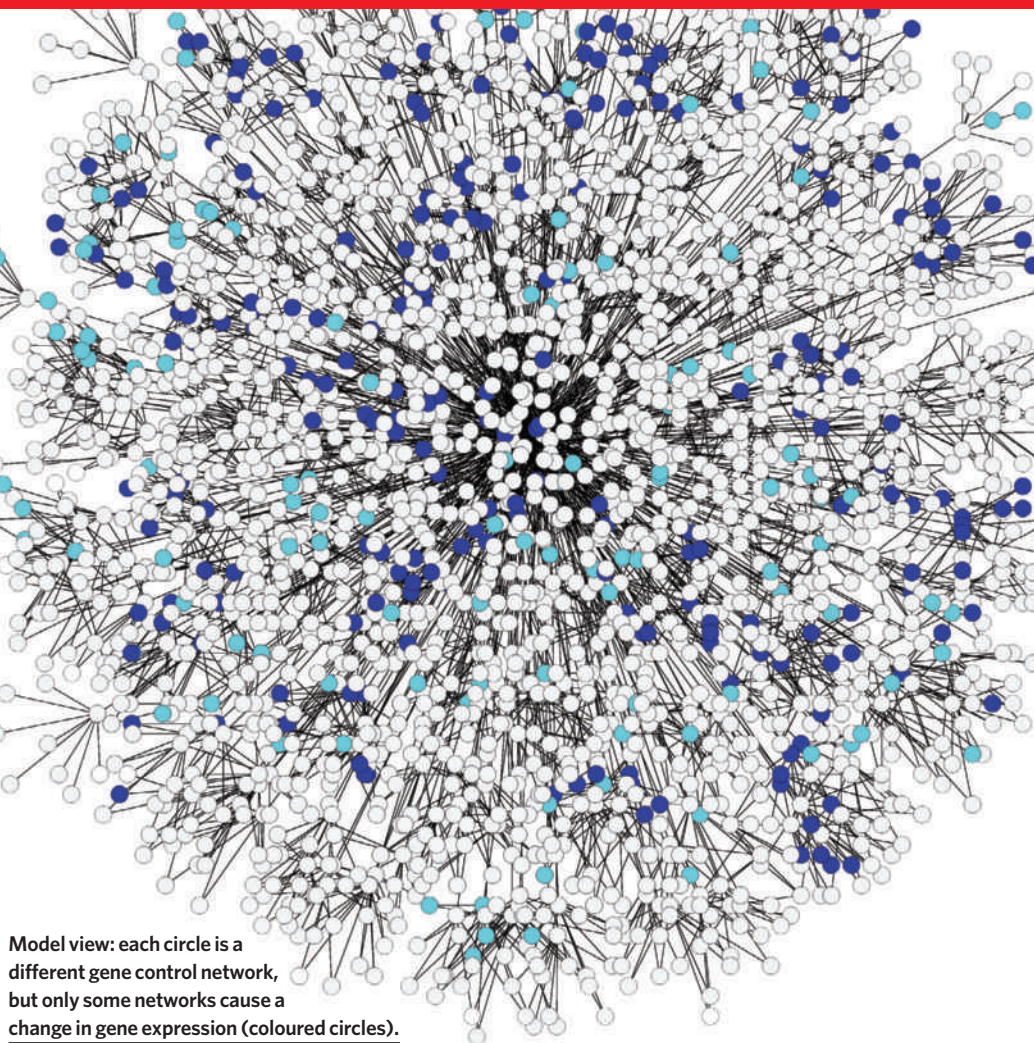
S. DARRAS

which the model starts, every gene has a given level of expression. The model then calculates what that level of expression, for each gene, will do to all the other genes, and makes the necessary changes. That gives a new pattern of expression — and again the model sees what that means in terms of regulating each gene, until the system settles down to a steady state — the end point of the development of the 'organism'. Each 'genotype', the set of instructions dictating which gene does what to which other gene, results in a different final state, or 'phenotype'. (Systems that never settle down are discarded as unviable.) This type of model was originally developed by the group of John Reinitz, of Stony Brook University, New York, and, despite its level of abstraction, it has described early *Drosophila* development well enough to predict mutant phenotypes.

Wagner and his colleagues then asked a question of their model: how many genotypes can generate the same phenotype? This is equivalent to asking how many different species' developmental programs can generate a perfect insect body plan. They explored this by allowing a given genotype to evolve by small steps (modifying one gene interaction at a time), but keeping only the genotypes that produced the same phenotype.

By repeating such 'neutral' evolutionary steps hundreds of times they could measure the 'neutral space' of genotypes that all produce the same phenotype — itself a fraction of the total space of all possible genotypes. Even with a relatively small number of genes, says Wagner, "the number of genotypes with any given phenotype is astronomically large". Not only large, but spread out; the team found that it was possible to get from their starting genotype to something completely different — to move from one end of that imaginary space to the other — while keeping the corresponding phenotype unchanged.

If the real world works like the model, regulatory networks can undergo dramatic reshuffles with no outward sign of what is going on at all. And Wagner's studies hint that being able to build the same edifice on such different — and changing — foundations may confer an evolutionary advantage. The team altered its algorithm to look at novel phenotypes that are just one mutation away from a starting genotype. So imagine genotypes in two far-flung parts of that imaginary neutral space, creating the same phenotype. Make a small mutation to one of those and it will produce a second phenotype; make a small change to the other and it will produce something very different. Wagner and his



Model view: each circle is a different gene control network, but only some networks cause a change in gene expression (coloured circles).

colleagues were not trying to get their model to innovate; but because the genotypes originally explored such a range of the possibilities without changing phenotype, they could later mutate in a very wide range of directions. So it may be that invisible variation in networks better prepares organisms for new challenges to come, by arming them with more possible solutions.

The real deal

This year, a group led by Mark Isalan, of the Center for Genomic Regulation in Barcelona, Spain, rewired a real genetic network⁶. Isalan picked out 22 DNA-binding regulatory proteins in *Escherichia coli* and then created close to 600 genetic sequences in which the DNA coding for a given protein was put under the regulatory control of another. They then introduced each one of these synthetic genes into normal bacteria. In essence, they added a handful of new links to *E. coli*'s endogenous regulatory network.

The group found that roughly 95% of the rewired bacteria survived: the bacterial regulatory network seems remarkably robust to arbitrary rewiring. And some of these engineered bacteria could out-compete the original bacteria in their ability to survive culture conditions such as 'heat-shock'. If genetic systems were like the computer code that humans build, this would be like dumping arbitrary

'Go-To' instructions in the middle of some air-traffic-control software and finding that not only did it cause very few crashes, but it also got more planes to their destination on time. The researchers concluded that, for *E. coli* at least, new links in regulatory networks are rarely a problem, and that they can even confer improved fitness in some environments.

But some biologists question the evolutionary significance of lab experiments, in which organisms are subject to less challenging conditions than in the wild. Systematic gene deletion in yeast, for example, has similarly shown that close to 80% of all protein-coding genes are not essential in normal lab culture. And yet their persistence as functional sequences in the genome is a strong indication that they contribute to the organism's fitness in the wild. So even if computer simulations or lab experiments suggest that network rewiring can be innocuous or even advantageous, the situation in real life might be different.

Come back to the yeast mating type: although the various genetic logics of *S. cerevisiae*, *C. albicans* and *K. lactis* seem to have identical outputs, in that they all generate two mating types, these may provide the various organisms with some unknown adaptive features that are of benefit in their different ecological niches in humans, breweries or cheese factories. "We don't know what hidden

S. CLIBERTI AND A. WAGNER

advantages if any might be associated with the various mating-type circuits found in those yeast species," Johnson says.

Sea-urchin development is one case in which alternative strategies for building the same adult body plan seem to have offered advantages in lifestyle. The two species in the genus *Heliocidaris*, for example, still share the shallow waters of southeastern Australia; their last common ancestor lived around the same time as the last common ancestor of chimps and humans. But they have dramatically different means of development. The *H. tuberculata* embryo develops first into a swimming and feeding larva called the 'pluteus', a form of indirect development that is shared by most of the 1,000 known species of sea urchins and probably represents the primitive mode for the clade. But *H. erythrogramma* develops direct from egg to adult in just a few days. Its entire developmental program has been completely scrambled to bring this about: its egg is 100 times bigger and the fates of the various cells that the egg divides into are completely different to those seen in *H. tuberculata*⁷. Even Wolpert's sacrosanct gastrulation is rearranged. It is as if, in a brief 4 million years, chimps had evolved to lay hard-shelled eggs while still developing into chimp-like creatures in the end.

Solving multiple problems

The sea urchins' developmental modes may represent equivalent solutions to one problem — how to build a spiny adult body that sticks to a rock — and yet be non-equivalent solutions to another — how to survive a youth in open water. Sea urchins with swimming and feeding larvae produce many small eggs, have broader geographic distributions, and show high mortality rates due to predators. Sea urchins that develop direct and fast produce fewer, bigger eggs and tend to stay put in remote niches; they have higher survival rates per embryo. In this case, changes that are 'neutral' with respect to

one aspect of an organism's phenotype seem to present advantages in others.

Many are unconvinced by these evo-devo studies. The common view that "the guiding hand of natural selection" is behind the evolution of all biological sophistication is greatly misleading, says evolutionary biologist Michael Lynch of Indiana University in Bloomington. Lynch thinks that most biologists underestimate the influence of chance, particularly in small populations in which neutral genetic changes can become the norm without any selective pressure by the process of 'genetic drift'. "Nothing in evolution makes sense except in the light of population genetics," Lynch says.

This is particularly true for gene regulatory networks, which can evolve by chance much faster than protein-coding genes because they can tolerate changes by point mutation, duplication or deletion with little effect on overall function. New binding sites for regulatory proteins, for example, can appear or disappear and the change can quickly become fixed in a population without loss of fitness⁸. Such constant reshuffling makes regulatory networks more complex than they need to be. It also creates redundancy — several parts of the network can perform the same task — and this allows further reshuffling to eventually bring about dramatic change in the way that things are run. Lynch thinks that Johnson's yeast studies are a particularly elegant demonstration of this.

Such "seemingly baroque architecture" of biological networks, as Lynch puts it, leaves developmental biologists with a problem: how to understand whole systems while relying on single-gene experiments. Traditionally, they have knocked out a gene in a model organism and drawn conclusions from that about the gene's function in a range of species in

which that gene's sequence is conserved. The evolutionary biologist Stephen Jay Gould long challenged this type of reductionist method, in which biological function of a gene is discussed 'all-other-things-being-equal' — the *ceteris paribus* paradigm⁹. Systems evo-devo makes it clear that all other things are not equal: the function of any gene cannot be defined outside its species-specific context.

Developmental biologists have to find a new way to describe the commonalities between organisms that were once ascribed to genes. The fact that tunicates and vertebrates both develop through a tadpole-like stage suggests that there must still be shared biological 'rules' governing the process. But if genes and regulatory links aren't conserved, then what exactly do they have in common? The shared features may actually lie in some other property of the regulatory networks.

It is this problem that Lemaire and others now

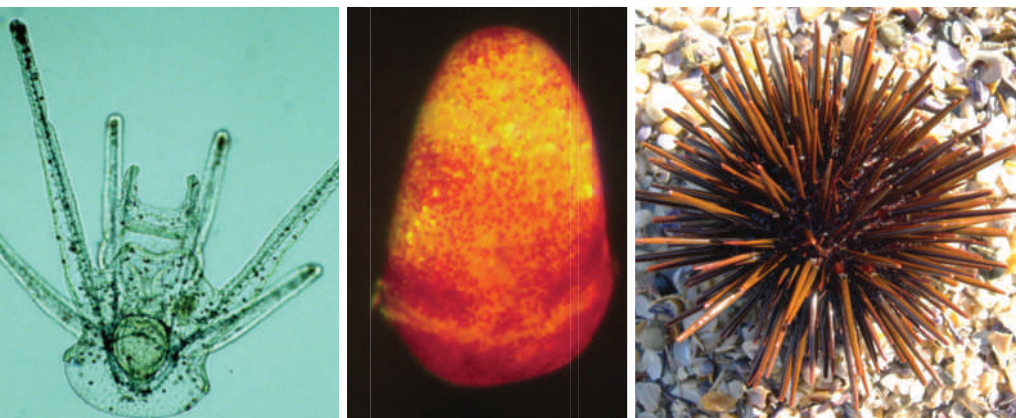
hope to address in tunicates. Because of the relatively simple body plan and the ease with which tunicate genes can be manipulated, researchers have come closer to building a complete regulatory network for tunicates than for any other model organism, one that shows how each regulatory gene interacts with every other, in every cell, and at every stage of development¹⁰. Lemaire is hoping to extract a 'network signature' that would characterize the process of gastrulation, perhaps a mathematical expression that describes the density of links between genes or the types of feedback loops used. Then researchers can try to see whether the same features are shared by gene regulatory networks in humans and other species that do rely on a Spemann organizer to gastrulate.

If Wolpert is right, and gastrulation is truly the most important event in your life, then that seems an answer worth knowing. ■

Tanguy Chouard is a senior editor at *Nature*.

"Nothing in evolution makes sense except in the light of population genetics."

— Michael Lynch



Heliocidaris sea urchins have very different larval stages (left and centre), yet reach a similar adult form.

1. Lemaire, P., Smith, W. C. & Nishida, H. *Curr. Biol.* **18**, R620–R631 (2008).
2. Schröder, R. *Nature* **422**, 621–625 (2003).
3. Lynch, J. A., Brent, A. E., Leaf, D. S., Pultz, M. A. & Desplan, C. *Nature* **439**, 728–732 (2006).
4. Tsong, A. E., Tsuchi, B. B., Li, H. & Johnson, A. D. *Nature* **443**, 415–420 (2006).
5. Ciliberti, S., Martin, O. C. & Wagner, A. *Proc. Natl Acad. Sci. USA* **104**, 13591–13596 (2007).
6. Isalan, M. *et al.* *Nature* **452**, 840–845 (2008).
7. Wray, G. A. & Raff, R. A. *Dev. Biol.* **132**, 458–470 (1989).
8. Lynch, M. *Nature Rev. Genet.* **8**, 803–813 (2007).
9. Gould, S. J. *The Structure of Evolutionary Theory* (Harvard University Press, 2003).
10. Imai, K. S., Levine, M., Satoh, N. & Satou, Y. *Science* **312**, 1183–1187 (2006).

See pages 281, and 295, and the Darwin special at www.nature.com/darwin.



M. NICHOLSON/CORBIS

AN EYE FOR THE EYE

“To suppose that the eye, with all its inimitable contrivances for adjusting the focus to different distances, for admitting different amounts of light, and for the correction of spherical and chromatic aberration, could have been formed by natural selection, seems, I freely confess, absurd in the highest possible degree ...

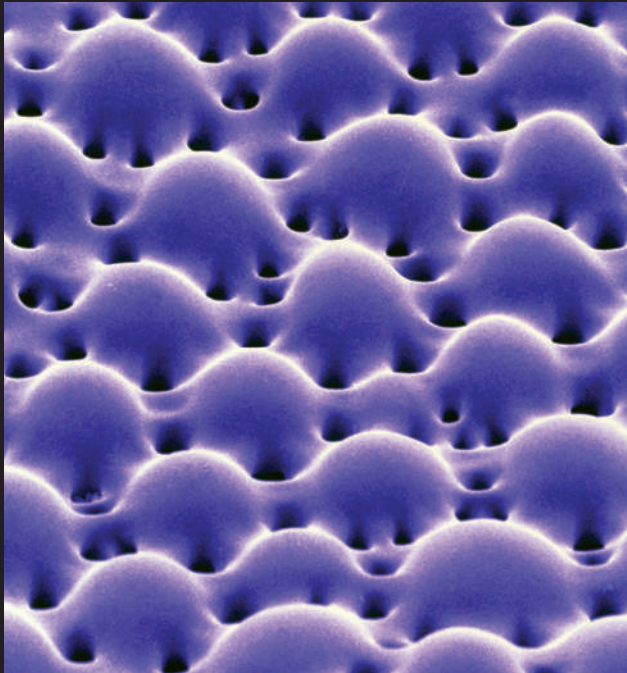
— Charles Darwin, *On the Origin of Species*



DRAGONFLY

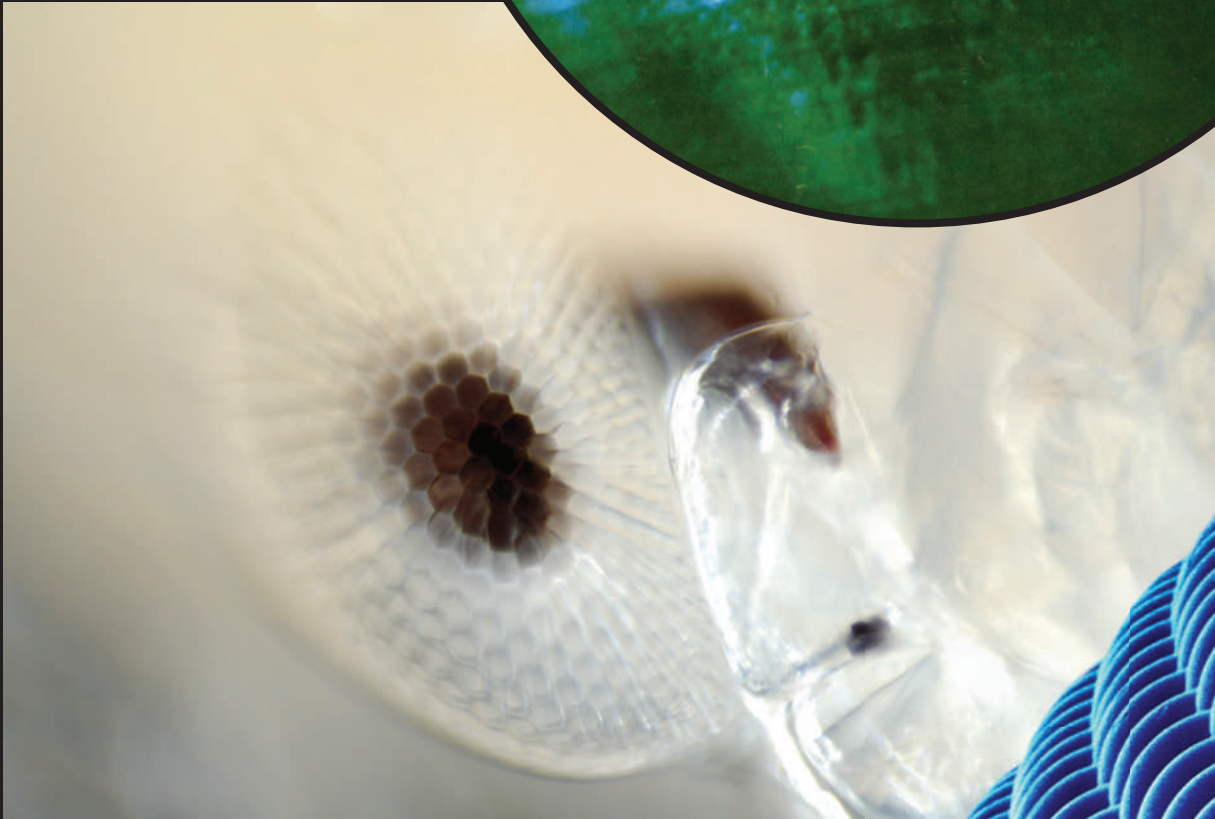
Eyes are largely built from building blocks designed for other things. The lenses of vertebrates use proteins that bacteria developed to deal with stress; the flexible guanine mirrors that make a cat's eye glow in the dark provide gas-proofing for the swim bladders of fish. And the components are put together in a bewildering number of ways. The brittlestar has an entire carapace pitted with optically tuned calcite crystals. The scallop eye uses a curved mirror to focus incoming light; the pram bug *Phronima* elongates each element of its eyes with a natural optical fibre. The dragonfly's compound eyes — each an array of up to 28,500 individual visual organs — may look ungainly, but they are extremely resistant to motion blur as it hunts on the wing.

BRITTLESTAR



LUCENT TECHNOLOGIES/BELL LABS/SPL

S. ZANK/NATUREPL.COM

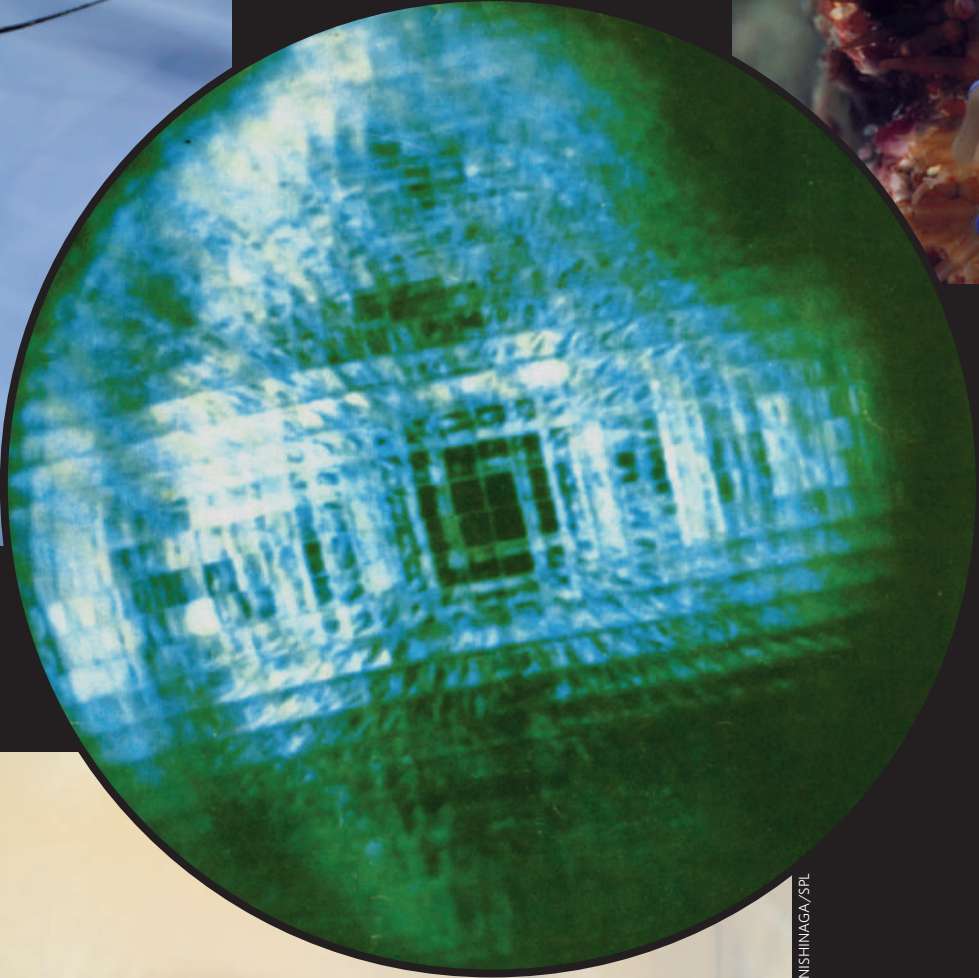


PRAM BUG

K. TAYLOR/NATUREPL.COM

SHRIMP

In dim conditions, an eye must gather what light it can, regardless of image clarity. By using box mirrors to focus light on a common point, the eyes of shrimp and lobsters enjoy more than 250 times the light-catching power of the human eye.



NATURE PRODUCTION/NATUREPL.COM



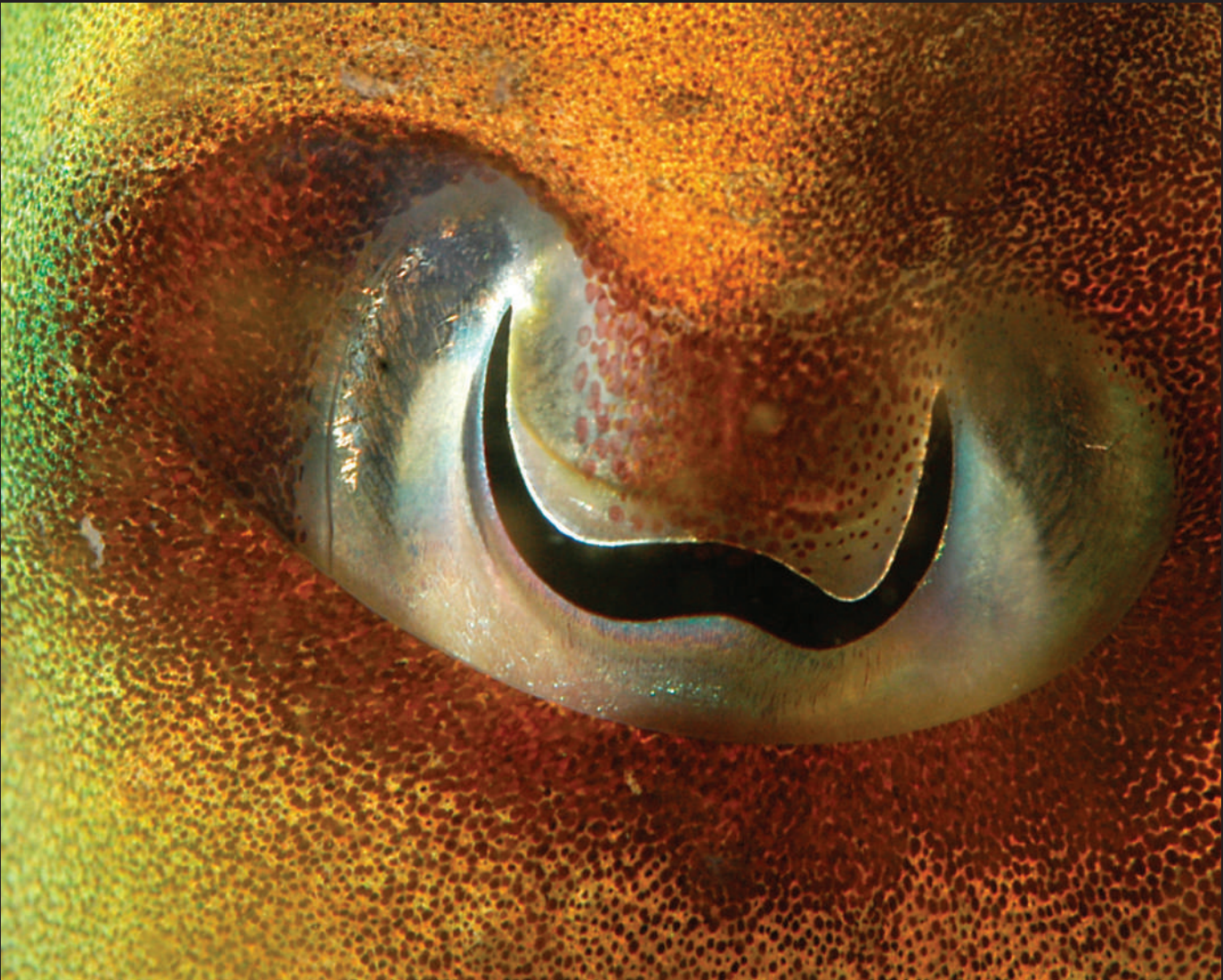
SCALLOP

HOVERFLY

The compound eye is a dead end in design terms; no matter how big you make it, it produces poor images. But it is well adapted to the zigzagging life of a hoverfly. The abrupt changes in the fly's direction of flight, which occur several times a second, help the eyes produce a fuller picture of the world.



SHRIMP: M. LAND; HOVERFLY: S. NISHINAGA/SPL



A. JAMES/NATUREPL.COM

Charles Darwin was well aware that the eye — so obviously, so brilliantly ‘designed’ — might represent an impediment to the acceptance of natural selection. In fact, it has come to be seen as one of evolution’s crowning glories. Across the animal kingdom, eyes have evolved in different ways, to different purposes, in exuberant diversity. Yet they are all sculpted by the lawfulness of light — and the imperatives of survival.

This panoply of eyes sees the world in different ways, some concerned with colour, others with movement, others with acuity. Yet time and again, unrelated eyes hit upon common solutions to the problem of how to safely gather and focus information from a sunlit world. The eyes of the cuttlefish (above) grow from invaginations of the skin; those of the human grow in part out of the front of the brain; each uses completely different receptors to pick up the light they focus. Yet cuttlefish and people see the world in the same way, through eyes whose similarities outweigh their deep differences — **Simon Ings.**



M. MOFFETT/MINDEN PICTURES/FLPA

SPIDER

The principal high-resolution eyes of true spiders are as big and as powerful as the eyes of small rodents. Subsidiary eyes are used to spot movement in the periphery, and sometimes, by harnessing the polarization of sunlight, to navigate.

MANTIS SHRIMP



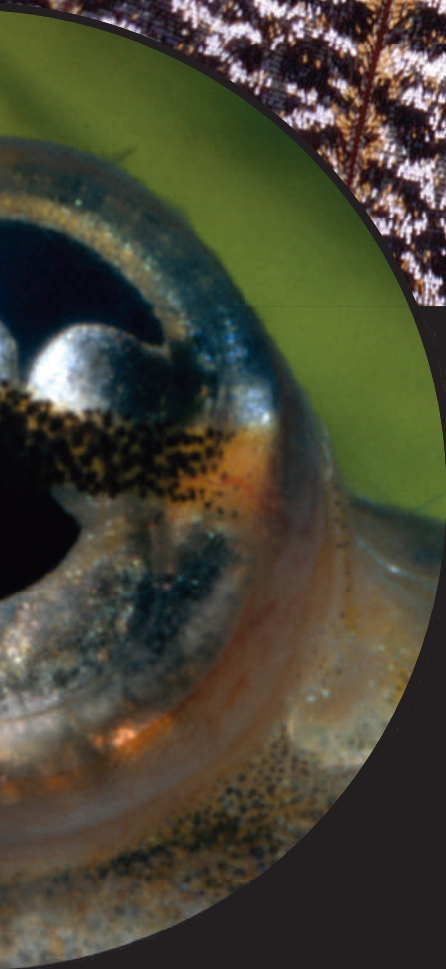
Human eyes are rare in their concern for producing images; most of the eyes in the animal kingdom are tuned to movement and colour. The band running across the eye of the mantis shrimp contains receptors tuned to 16 different wavelengths (humans typically have three), giving colour vision of extraordinary subtlety and complexity. Nor are two eyes a universal norm. The fish *Anableps anableps* effectively has four eyes in two sockets; each eye has one half for seeing above water and one half for seeing below. Another fish, *Bathylchnops exilis*, has two pairs, one set to look up, which is how it sees its prey, and one to look down — but for what, nobody knows. The better your predators can see, the more need there is for disguise. For some animals, looking inedible is not enough: the unnervingly detailed mock eye on a moth wing is meant to convince predators that something big is looking right back at them.

MOTH: ALAMY; ANABLEPS: P. ZAHL/NATL GEOGRAPHIC/GETTY IMAGES

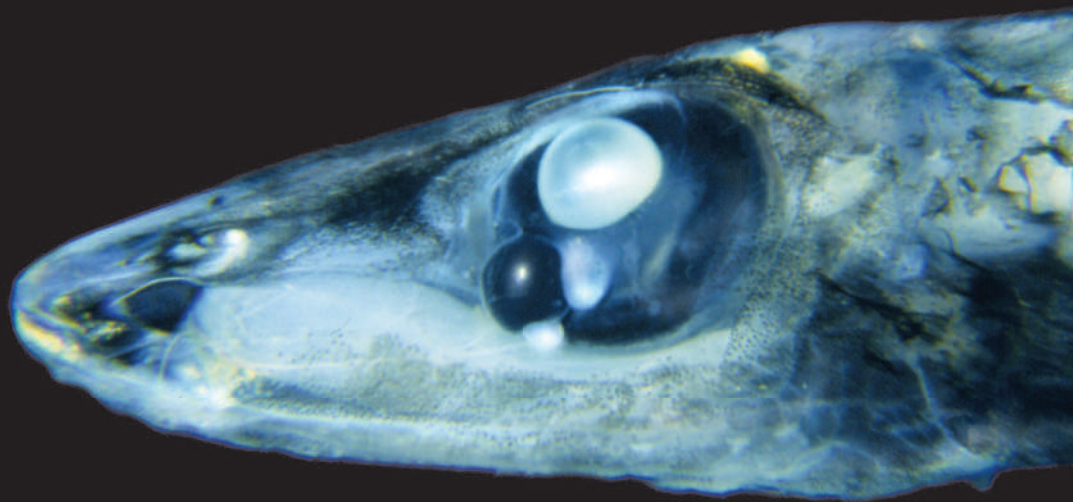
MOTH



ANABLEPS

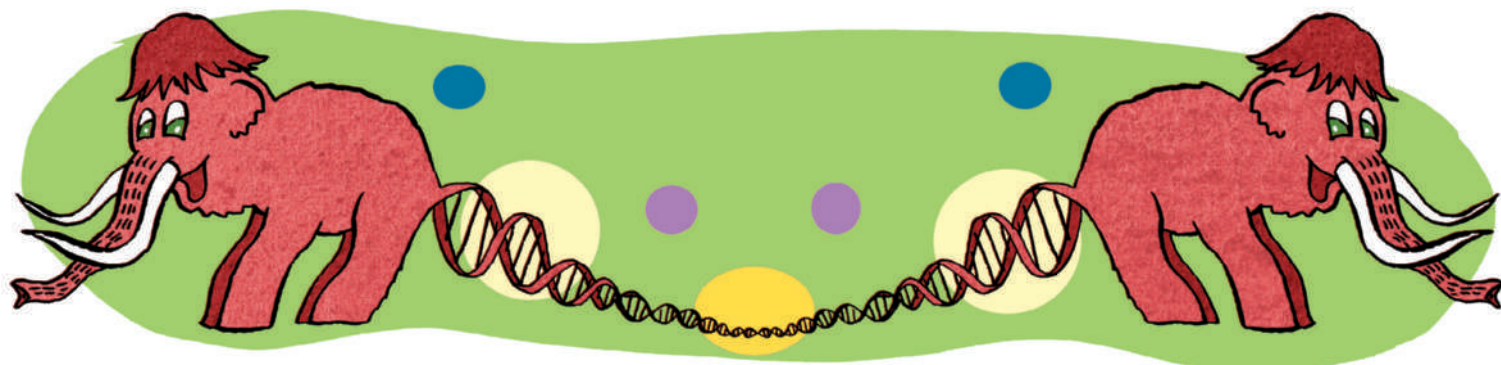


BATHYLYCHNOPS



... Yet reason tells me, that if numerous gradations from a perfect and complex eye to one very imperfect and simple, each grade being useful to its possessor, can be shown to exist; if further, the eye does vary ever so slightly, and the variations be inherited, which is certainly the case; and if any variation or modification in the organ be ever useful to an animal under changing conditions of life, then the difficulty of believing that a perfect and complex eye could be formed by natural selection, though insuperable by our imagination, can hardly be considered real.”

The order of the reference list and citations have been corrected in this PDF.



LET'S MAKE A MAMMOTH

Evolution assumes that extinction is forever. Maybe not. **Henry Nicholls** asks what it would take to bring the woolly mammoth back from the dead.

In 1990 the late Michael Crichton gave the idea of reviving extinct species a slickly plausible and enormously entertaining workout in his novel *Jurassic Park*. At that time the longest genome that had ever been sequenced was that of a virus. The best part of 20 years on, hundreds of animal genome sequences have been published. This week, for the first time, the genome of something undoubtedly charismatic and definitively extinct joins the list: the mammoth (*Mammuthus primigenius*)¹. If you want to bring a species back to life, the mammoth would be almost as dramatic as a dinosaur. And unlike *Tyrannosaurus rex*, the mammoth has close living relatives to lend a hand.

It is a fair bet that a complete genome and closely related species would make it easier to pull a Crichton on a mammoth than on a dinosaur. But easier is far from easy. To put flesh on the bones of the draft sequence — to go from a genome to a living, breathing beast — would require you to master, at the very least, the following steps: defining exactly the sequence or sequences you want for your creatures; synthesizing a full set of chromosomes from these sequences; engulfing them in a nuclear envelope; transferring that nucleus into an egg that would support it; and getting that egg into a womb that would carry it to term. None of those steps is currently possible. From getting a definitive sequence to harvesting eggs from an elephant there are all-but-insurmountable obstacles at every stage,

and no evidence that anyone is going to work very hard to solve them. But do any of them actually make the dream impossible?

The sequence

The first step in this mammoth challenge is to obtain a sequence good enough for us to contemplate using it as the basis for a living being. The sequencing of long-dead DNA such as that of mammoths uses fragments at various levels of degradation. To detect and correct the base-pair changes that can occur after death and to avoid the inevitable errors involved in assembling millions of these tiny fragments into a coherent sequence, it is necessary to sequence much more than a single genome's worth of DNA. "If we want fewer than 1 error in 10,000 base pairs — a reasonable quality genome — we would need to sample in the order of 12-fold coverage," says Svante Pääbo, director of the genetics department at the Max Planck Institute for Evolutionary Anthropology in Leipzig, Germany, who has worked on Neanderthal genomes². The genome published today has roughly 0.7-fold coverage. 'Reasonable quality' for science does not mean the sort of genome you would want to live with: in a human genome that error rate would mean 300,000 mutations.

Coverage can be improved as long as there's the money to do it, but old samples offer particular challenges: a lot of contamination by bacterial, fungal and other species' DNA.



Darwin200

Thirty-five-fold coverage, which Pääbo says is as good as it gets, would be "extremely costly and extremely time-consuming", according to Eske Willerslev, head of the Ancient DNA and Evolution Group at the University of Copenhagen. Even cheaper sequencing, however, and the possibility of better preserved and prepared samples, mean that those expenses of cost and time will shrink. Willerslev sees nothing to stop researchers from producing a mammoth genome as good as any genome today at some point in the future. Whether such a genome would be good enough for a living being remains a somewhat open question — but with time and effort, it's plausible that a sufficiently error-free genome can be arrived at.

A sequence on its own, though, is not enough: researchers will need to work out exactly how it divides up into chromosomes. The obvious solution would be to tot up the number of chromosomes in an intact mammoth cell and sift through the genomic data looking for their beginnings and endings. But even the very best mammoth material falls short of this kind of preservation (see 'You need to do more than thaw'). "We have no idea — yet — how many chromosomes mammoths had," says Hendrik Poinar, a geneticist at McMaster University in Ontario, Canada. Kerstin Lindblad-Toh, co-director of the genome sequencing and analysis programme at the Broad Institute in Cambridge, Massachusetts, says that the institute will release a sequence of the African elephant (*Loxodonta africana*) to seven-fold coverage some time in 2009. When they do, the mammoth geneticists

"Genome synthesis is further developed today than genome sequencing was when Crichton wrote *Jurassic Park*."

ILLUSTRATIONS BY: S. KAMBAYASHI

will be all over it. But it will take an immense amount of work to identify and locate all the chromosome changes, gene deletions, duplications and rearrangements that mammoths will have acquired since they diverged from their African ancestors 7.6 million years ago. A sequence for the Indian elephant (*Elephas maximus indicus*), which is more closely related to the mammoth, would be of further help.

One chromosome offers particular problems. In mammals the Y chromosome is typically very small and hard to sort out, in part because it is remarkably repetitive. Researchers have sidestepped the issue in the elephant genome by sequencing a female. "The X chromosome is hard enough to assemble and the Y chromosome is the hardest chromosome out there," says Lindblad-Toh. The dinosaurs in *Jurassic Park* were designed to all be female, to avoid unwanted breeding; cloned mammoths might all be females, too, at least for the first generation, just because it would be easier.

There are other repetitive regions that will be hard to sequence with confidence, most notably the centromeres, which help chromosomes to get where they are meant to go in cell division. It is almost impossible to work out the centromeres' exact sequence, says Bill Earnshaw of the Wellcome Trust Centre for Cell Biology at the University of Edinburgh, UK. "You just get hopelessly lost," he says. "It's like being in a forest where all the trees look identical."

But this need not be a sticking point, as artifice can make do for accuracy. Just this year Earnshaw and his colleagues created a human artificial chromosome that contained a synthetic but fully functional centromere³. In principle, that could work for a synthetic mammoth chromosome, says Earnshaw. Replacements for the shorter repetitive sequences at the end of chromosomes, called telomeres, are also doable. And although the sequences will need specific sites at which chromosome replication can start, too, Earnshaw says that "any long enough strand of DNA will have sequences that can function as origins of replication".

Finally, there is the question of genetic variation. Most published mammalian genomes — the mammoth draft included — provide only a single version of all the genes and other sequences in the genome. But mammals have two versions of each gene — one from their mother and one from their father. Building a mammoth with chromosome pairs in which the two chromosomes were identical would be a recipe for trouble, as the effects of any bad gene would be felt to their fullest.

You need to do more than thaw

Making a nucleus will be difficult; why bother if some are already lying around? Some argue that finding frozen mammoth cells that have well-preserved nuclei is the obvious starting point for making a mammoth.

In Japan, scientists¹² have recently succeeded in cloning mice from cells that had been frozen for 16 years. This led them to venture that "nuclear-transfer techniques could be used to 'resurrect' animals ... from tissues frozen for prolonged periods without any cryopreservation". They mentioned mammoths

specifically in this context. But those who work with ancient DNA on a daily basis are unconvinced. Sixteen years in a -20°C laboratory freezer is a poor substitute for ten millennia or more lying in permafrost — and who knows how long unfrozen before that?

Eske Willerslev, head of the Ancient DNA and Evolution Group at the University of Copenhagen, points out that it takes time, and repeated cycles of tissue-damaging freezing and thawing, before a dead animal gets incorporated into the permafrost. That

gives plenty of opportunity for DNA to degrade. What's more, Willerslev says, many microbes are active even at subzero temperatures, causing the samples to degrade further. All this, he thinks, explains why he has never seen or heard of a mammoth cell with a nucleus that's even intact, let alone transferable. Hendrik Poinar of McMaster University in Ontario, Canada, puts the problem with frozen-tissue approaches more succinctly: "I am currently unaware of -20° freezers that date to 20,000 years." **H.N.**

Identifying different versions of genes would add yet more to the sequencers' to-do list, but it would be crucial to success.

DNA synthesis

With the genome sequenced in painstaking detail, glitches corrected, chromosomes identified, key repeat sequences written in appropriately and genetic variation introduced, it's time to turn towards the DNA synthesis itself. The largest totally synthetic genome produced so far is that of the bacterium *Mycoplasma genitalium*⁴. This contained 582,970 base pairs;

the mammoth weighs in at 4.7 billion, half as many again as are found in the human genome. Assuming mammoths turn out to have the same number of chromosomes as elephants, the task would be broken down into making 56 separate chromosomes, each an average of some 160 million base pairs long.

Short strings of bases that are made in the test tube can be assembled into respectable double-stranded stretches of DNA about 8,000 base pairs long without too much error, says Ralph Baric, a microbiologist at the Carolina Vaccine Institute in Chapel Hill, North Carolina; a range of companies will synthesise

such sequences for less than a dollar a base pair, and the reagents cost much less. But as neighbouring stretches are joined together *in vitro*, the DNA molecules become increasingly unstable. The team that put together the *M. genitalium* genome at the J. Craig Venter Institute in Rockville, Maryland, dealt with this by inserting the unstable chunks of DNA into 'bacterial artificial chromosomes'. The various components could then be stitched together in the relatively welcoming environment of *Escherichia coli*. For the last steps, the researchers took the largest pieces assembled in *E. coli* and inserted them into yeast artificial chromosomes, which are larger. These were then recombined in living yeast cells to produce constructs that had the entire genome in them.

This approach is impressive in



terms of speed and cost, says Drew Endy of the Department of Bioengineering at Stanford University in California. But it's unlikely to be scaled up to accommodate mammoth-sized chromosomes in any straightforward way. After all, he says, at just over 12 million base pairs, the entire yeast genome is much smaller than a medium-sized elephant chromosome: "I would wonder if yeast could handle so much exogenous DNA." This concern is echoed by Baric: "The bigger it is, the quicker there are going to be nicks or breaks, allowing for degradation or deletions in essential genes."

Even if it became possible to synthesize a stable mammoth chromosome in sufficient quantities, this would then have to be repeated for all the other chromosomes. "I don't think we're going to be seeing any mammoths any time soon," concludes Baric. But it will not be long before synthetic biologists develop a certain confidence in synthesizing microbial genomes from scratch, he says. The technologies that are used to do so will, he predicts, give a good indication of whether it might one day be possible to reconstruct the genome of a large mammal such as a mammoth. It is worth remembering that genome synthesis is further developed today, in terms of the maximum lengths achieved, than genome sequencing was when Crichton wrote *Jurassic Park*. And look how sequencing has progressed since then.

Wrapping it up

Once the chromosomes have been synthesized, they need to be put into a nucleus. Cells wrap chromosomes back up into nuclei whenever they divide, so this part of the process has been fairly well studied. In the 1980s, researchers

discovered that naked DNA added to extracts from the eggs of frogs quickly becomes wrapped up in proteins that condense it into chromatin; then membrane fragments bind and fuse to form a functional nuclear envelope around the chromatin. "Such artificial nuclei are capable of DNA replication and some DNA transcription," says Douglass Forbes, professor of cell and developmental biology at the University of California, San Diego. Forbes thinks that anyone trying to make a mammoth nucleus in the foreseeable future would be best advised to stick with frog extracts, rather than use fragments of the nucleus from some more similar creature such as an elephant. "Mammalian eggs, which are fertilized internally, might have much less ability to assemble nuclei," she says. When later transferred into the cytoplasm of an elephant cell, elephant nuclear proteins might complement or replace the frog-specific counterparts used to make the mammoth pseudo-nucleus, she says, giving it a more properly mammalian look and feel.

It would also be a challenge to ensure that the nuclear membrane engulfs all the mammoth chromosomes. "You will somehow need to keep them together when you inject them," says Eric Schirmer at the Wellcome Trust Centre. If you don't, miniature nuclei may form that contain a random rabble of chromosomes. The way these chromosomes sit with respect to one another might also affect gene expression; how to engineer the correct configuration, nobody knows.

Egg collection

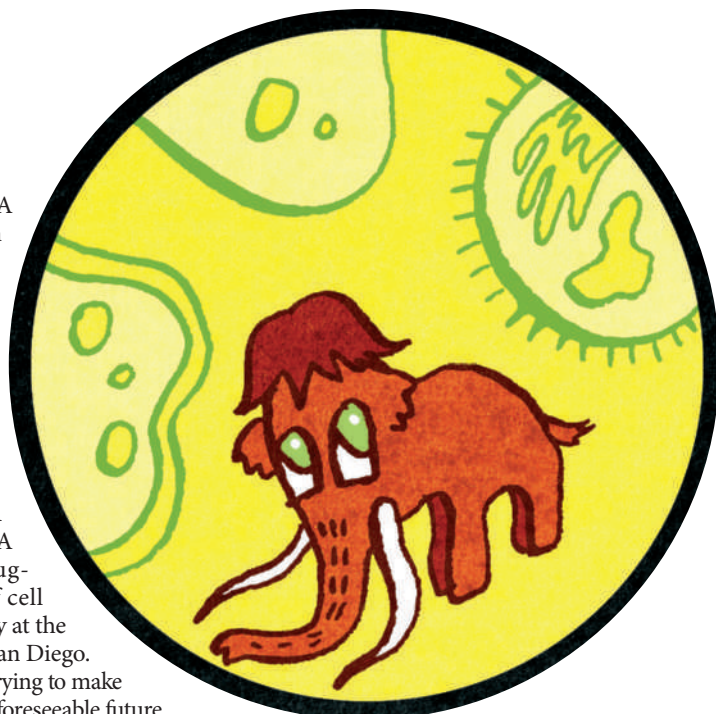
It's almost time to contemplate the vagaries of nuclear transfer, but not before you have sourced your elephant eggs, and these are likely to be in pretty short supply. Female elephants ovulate on a 16-week cycle, although they regularly skip five or so years owing to gestation and lactation. Stopping these natural breaks in cycling would be both cruel and unproductive; cow elephants that don't gestate have a strong tendency to develop massive ovarian tumours. When they do ovulate, only one oocyte is released; a litter of little elephants would be a death sentence, and even

twins are remarkably rare.

On the positive side, though, elephants' infrequent ovulations are preceded by an early warning; uniquely among mammals that have been studied to date, elephant ovulation involves not one but two surges in luteinizing hormone, separated by 18–20 days. "The first hormone peak dissolves the vaginal mucous and forwards just one follicle for development," says Thomas Hildebrandt of the Institute for Zoo and Wildlife Research in Berlin, Germany. "The second peak stimulates ovulation."

In other creatures it would be quite straightforward to get the follicle in which an egg is developing out of the ovary after this surge of harbinger hormones; you use ultrasound to guide a harvesting implement up the reproductive tract, or perform a laparoscopy, during which the abdominal cavity is inflated to make room for the job to be done surgically.

Unfortunately, quirks of elephant biology rule out both these approaches. Whereas the entrance of the vagina is pretty simple to locate in most mammals, elephants have more than a metre of urogenital canal between the outside world and the hymen. This canal is as far as a bull-elephant's penis gets; the hymen remains intact even after copulation, which may be an evolutionary hangover from the elephant's aquatic past. Hildebrandt and his colleagues have developed a way to navigate an instrument up the canal, through the tiny aperture in the hymen that lets sperm in, along the vagina and into the uterus; they use it to perform artificial insemination with sex-selected sperm. But Hildebrandt says that even with such instruments threaded into the womb it would be almost impossible to locate a single mature follicle without some extra guidance, and the ovaries are too deep inside the abdominal cavity for the precise position of the follicle to be



Let's fake a mammoth

The complete genome sequence of the African elephant (*Loxodonta africana*), expected soon from the Broad Institute in Cambridge, Massachusetts, should be a great help in identifying regions of the mammoth genome that vary from those of the elephant and starting the search for the key genetic differences that are unique to mammoths. But it might also provide a more direct

and plausible way to make what one might call pseudo-mammoths.

With some idea of the genes involved in hair formation, coat colour and tusk development, for example, it might be possible to use genetic engineering and selective breeding to produce an elephant that looks remarkably mammoth-like, says Svante Pääbo of the Max Planck Institute for Evolutionary Anthropology

in Leipzig, Germany. It wouldn't be a mammoth in a way that satisfied a purist, or an ecologist or a dreamer set on restoring something magnificent and lost to the world. But it could perhaps be good enough for, say, amusement-park work; and it would avoid some of the trickiest synthetic biology. "That's the furthest I could imagine I would see in my lifetime," says 53-year-old Pääbo.

H.N.

native elephant nucleus to make space for the synthetic mammoth nucleus. But that might not be all the preparation needed for the nuclear transfer. The mitochondria — organelles that provide cells with energy through respiration — have their own genomes, specific to each species. Hybrids with a nucleus from one species and mitochondria from another can be viable — an embryo with sheep mitochondria but nuclear DNA from a mouflon developed into an animal, called Ombretta, in 2001 (ref. 7) — but there are risks of incompatibility. The close evolutionary relationship between elephants and mammoths reduces the chances of incompatibility, but Stefan Hiendleder, professor of animal science at the University of Adelaide in Australia, cautions that there would still be risks: human cells that have had their mitochondria replaced with those from other primates have failed to show proper respiration⁸. This means that researchers might need to rustle up some synthetic mammoth mitochondria and insert them into the elephant cytoplasm. With the mammoth mitochondrial genome already well sequenced⁹, this should be relatively simple. And all the elephant mitochondria would have to be cleaned out, in case they hybridize with the newcomers, warns Hiendleder.

Nuclear transfer remains a fickle and inefficient way of producing new mammals even without such concerns. Only a few of the transfers result in embryos and not all of those embryos manage to establish a placenta. Of those that do, many abort spontaneously and the few successful live births frequently have developmental abnormalities.

But it is reasonable to hope that things may get better, says Xiuchun Tian of the Center for Regenerative Biology and Department of Animal Science at the University of Connecticut in Storrs. The most likely explanation for the inefficiency of nuclear transfer is that the genes are not being expressed in a manner appropriate for embryonic development. This, says Tian, is probably down to errant 'epigenetic' signals — inherited patterns of DNA methylation, histone modification, micro-RNA presence and chromatin structure, all of which can have a crucial effect on gene

revealed by ultrasound. Laparoscopic surgery is also out of the question, as elephants have no pleural space between their body wall and lungs. "Inflating the abdominal cavity during laparoscopy would compress the lungs and kill the animal," says Hildebrandt.

There may, however, be an ingenious way out of this bind. Cryobiologists have in the past transplanted tissue from the ovaries of one animal into the ovaries of another. The procedure has worked even if the tissue had been frozen and thawed out between leaving the donor and being grafted into the recipient. Not only has this work led to the treatment of infertility after chemotherapy, it has also raised the possibility that ovarian tissue from endangered species could be transplanted into laboratory animals, making them a source of eggs-on-demand. For Hildebrandt, the only realistic way of getting a steady source of elephant eggs would be to collect tissue from the ovary of a recently deceased elephant and graft slivers of it into a laboratory mouse or rat with a suppressed immune system that won't reject them.

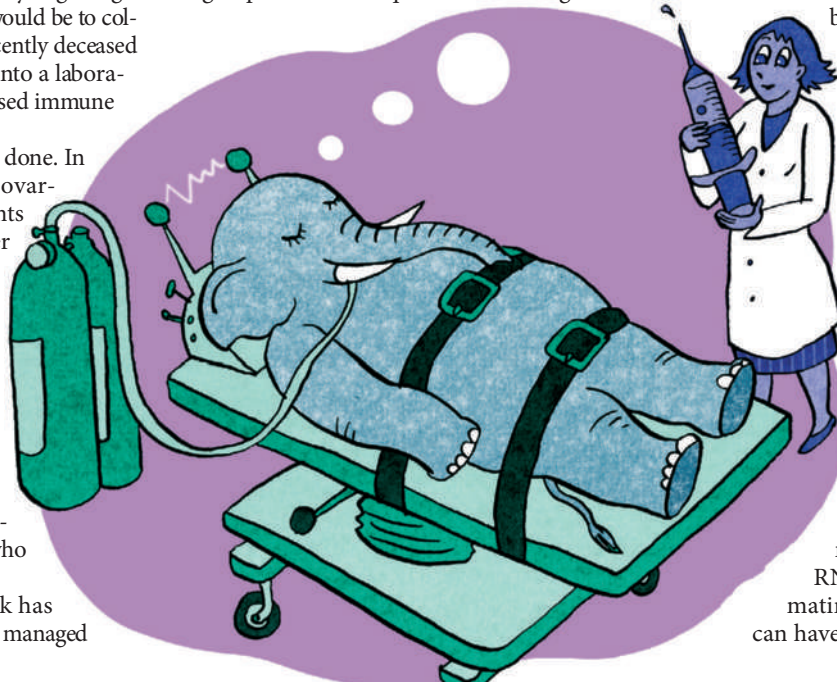
This has in fact already been done. In the 1990s, frozen samples of ovarian tissue from African elephants culled in South Africa's Kruger National Park were thawed and transplanted into mice, and afterwards researchers saw what seemed to be mature follicles develop⁵. "But we didn't have enough material to assess whether those eggs were competent," recalls John Critser, now professor of comparative medicine at the University of Missouri in Columbia, who led the study.

Although the elephant work has gone no further, researchers have managed

to produce live and apparently healthy offspring from eggs that have come from tissue transplanted into another species. Such successes, though, have been with closely related species — rats acting as hosts for mouse tissue⁶, for example. It's harder to see how elephant eggs would mature successfully in a mouse, not least because the mouse's oestrus cycle lasts just 4–6 days. It would probably be necessary to remove the mouse's pituitary gland and subject it to a hormonal cycle that approximates an elephant's with hormone treatments, says Critser. "It's not something that you could use easily right now for the production of oocytes. But that's not to say that it couldn't be developed."

Nuclear transfer

With a ready supply of eggs, it should not take long to perfect techniques for removing the





elephant embryo transfer, Hildebrandt has thought about what it would take. First, he recommends inserting an arm up the urogenital canal to inject some sperm-free elephant ejaculate in the direction of the hymen on the day that hormones reveal the elephant is ovulating. "We think the sperm-free component is needed to prepare the female's uterus to receive an embryo," he says. After that, the transfer of the cloned mammoth embryo into

the uterus, a total distance of some 2.5 metres, should be possible using the same apparatus used for artificial insemination, says Hildebrandt. "We are coming quite close to the oviduct, which would be the place to put a cloned embryo back into an elephant." An embryo at the four-cell stage would need to be transferred about two days about ovulation, he says.

At that point, the last concern becomes whether a mammoth fetus would be suited to the uterus of its surrogate mother. Evidence from preserved mammoths suggests that size, at least, should not be a problem. An Indian elephant calf can weigh around 120 kilograms at term and stands around 1 metre tall. Dima — a famous mammoth calf unearthed in northeastern Siberia in 1977 — is

"The urogenital canal is as far as a bull elephant's penis gets."

estimated to have been about the same mass and height when he died aged 7–8 months. It's a similar story for Lyuba, a calf discovered in Russia's Yamal Peninsula last year. Although Lyuba's mass when she died has not yet been definitively ascertained, initial reports suggest she was only 90 centimetres tall and 4 months old, says Daniel Fisher, professor of ecology and evolutionary biology at the University of Michigan in Ann Arbor and one of an international team poring over her spectacularly preserved remains. Most evidence indicates that newborn woolly mammoths were about the same size as newborn elephants, he says. "They could even have been smaller." Mercifully, we probably need not concern ourselves with how to incubate a preterm mammoth fetus.

Birth and after

A single artificial mammoth would be a freak, not a species; once she was born others — including, ideally, males — would have to follow. Their genomes would have to be tweaked to ensure a certain diversity. A place for them to live would have to be found, as would an ecosystem into which they could integrate. It would be a huge undertaking — just as synthesizing mammoth chromosomes and reprogramming them into embryo-friendly nuclei would be. Perhaps the whole idea will remain too strange, too expensive, too impractical, even too unappealing for anyone to take seriously. But the fact that just 15 years ago cloning mammals was confidently ruled out by many as being impractical should give people pause before saying any such thing is impossible. On Darwin's 200th birthday in 2009, reoriginating extinct animal species will still be a fantasy. By 2059, who knows what may have returned, rebooted, to walk the Earth?

Henry Nicholls is a science writer who lives in Greenwich, England. His most recent book is *Lonesome George*.

1. Miller, W. et al. *Nature* **456**, 387–390 (2008).
2. Green, R. E. et al. *Cell* **134**, 416–426 (2008).
3. Nakano, M. et al. *Dev. Cell* **14**, 507–522 (2006).
4. Gibson, D. G. et al. *Science* **319**, 1215–1220 (2008).
5. Gunasena, K. T. et al. *Anim. Reprod. Sci.* **53**, 265–275 (1998).
6. Snow, M., Cox, S.-L., Jenkin, G., Trounson, A. & Shaw, J. *Science* **297**, 2227 (2002).
7. Loi, P. et al. *Nature Biotechnol.* **19**, 962–964 (2001).
8. Hiendler, S., Zakhartchenko, V. & Wolf, E. *Reprod. Fert. Dev.* **17**, 69–83 (2005).
9. Gilbert, M. T. et al. *Science* **317**, 1927–1930 (2007).
10. Yang, X. et al. *Nature Genet.* **39**, 295–302 (2007).
11. Takeuchi, Y., Yoshizaki, G. & Takeuchi, T. *Nature* **430**, 629–630 (2004).
12. Wakayama, S. et al. *Proc. Natl Acad. Sci. USA* Advance online publication doi:10.1073/pnas.0806166105 (2008).

See also pages 281, 295, 330 and 387, and online at www.nature.com/darwin.

expression¹⁰. The recent discovery that the nuclei of normal body cells can be induced into an embryonic-like state without the need for nuclear transfer will help researchers to understand the effect of these epigenetic signals and how to manipulate them, she says.

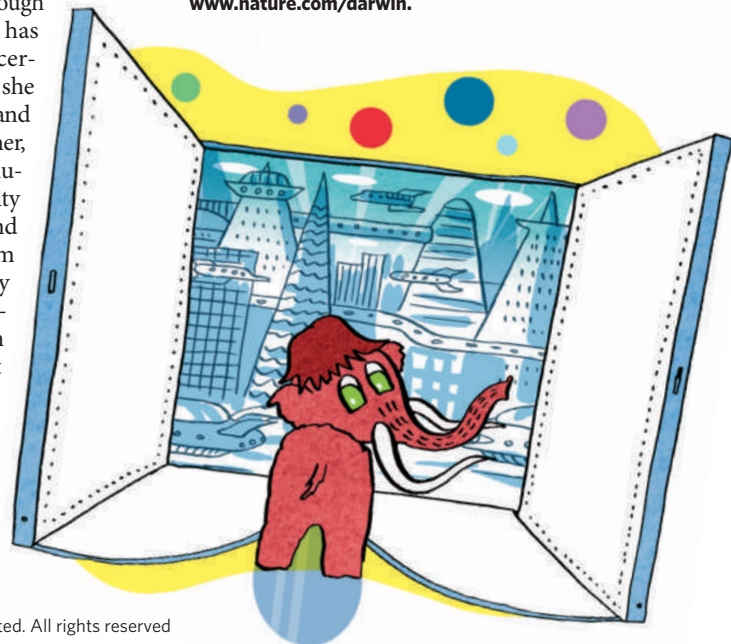
Better still, there may be an alternative way to dress up a synthetic nucleus in suitably epigenetic garb. Even if an early embryo is doomed not to go all the way to term, it can still be used as a source of stem cells. These could then be introduced into normal elephant embryos to create chimaeras — in which some cells are mammoth and some elephant. Such chimaeras may stand a better chance of developing to term, and although they wouldn't be mammoths, they would be a way by which mammoths might then be made.

If enough chimaeras are created in this way, some should end up with mammoth cells in their ovaries or testes, giving you elephants that produce mammoth eggs or sperm. 'Germ-line chimaeras' of this type have already been created in several species. In 2004, Japanese scientists¹¹ made salmon with trout cells in the testes that produced sperm capable of fertilizing trout eggs and producing bona fide baby trout.

Gametes from germline chimaeras would be much more likely to be properly imprinted, says Tian — indeed, once you have a supply of mammoth eggs and mammoth sperm you might well consider the project close to bearing fruit.

Embryo transfer

With a fertilized mammoth egg — either made through direct nuclear transfer or through the *in vitro* fertilization of mammoth eggs with mammoth sperm from chimaeras — the mammoth challenge comes to its final stage. Although the inaccessibility of elephant eggs means that nobody has ever performed



CORRESPONDENCE

Key discoveries often originate with lone researchers

SIR — Your News Feature 'Group theory' (*Nature* **455**, 720–723; 2008) highlights the dwindling contribution of single authors in the face of the productive increase in collaboration among scientists. Sociologist Brian Uzzi is not alone in questioning whether a lone researcher can still capture the moment when lightning strikes.

Uzzi uses some general metrics to suggest a correlation between the number of citations and the number of authors per paper. But before we rush to embrace these heuristics as a basis for designating best practices for collaboration, our social scientists need a better understanding of how high-impact discoveries in the life sciences come about, for example at different stages of the drug-discovery process.

Moving along this continuum from early discovery, through the translational stage and on to clinical trials, we would expect to see increases in the aggregate levels of funding and in the number of authors and papers published per year. Consequently, epidemiological evaluation will produce metrics and heuristics that more heavily reflect the practices employed in those later stages of the discovery process.

But social collaboration patterns probably vary with each stage. In the discovery phase, for example, projects are often driven by sole investigators, with support from a tightly knit group of colleagues from universities and biopharmaceutical enterprises, frequently generalists who move readily across disciplines and are at ease with counterintuitive interpretations. These small groups act as the initial nucleating agent, but they are then radically transformed into a complex web of interdependent

specialists as the drug, vaccine or biological candidate is developed and moves towards the clinic.

With the explosion of available information, publishers and granting organizations are understandably desperate, seeking tools that more effectively predict high-impact science. But in searching for these tools through the lens of social network analysis, we must not lose sight of the key contribution by a project's pioneer. We shall then see more clearly how collaboration patterns alter as an area of scientific discovery matures.

Readers, reviewers and grant administrators should not be biased against early-stage papers with just one or two authors. Lightning can still strike the solitary explorer whose mind is prepared.

Shawn J. Green, Jon Brendsel
PubScholar, 11212 Sorrel Ridge Lane,
Oakton, Virginia 22124, USA
e-mail: shawn@pubscholar.com

Significant confusion in scientists' grasp of statistics

SIR — I agree with the points you make about statistical significance under the heading 'Significant' in your News Feature 'Disputed definitions' (*Nature* **455**, 1023–1028; 2008). However, you do imply that the term 'significant' means simply above or below the 5% level — a figure chosen by the statistician R. A. Fisher for practical reasons and used in the days when people did arithmetic by hand and referred to printed tables.

Nowadays, of course, personal computers do more general calculations and report probability (*P*) values directly. A *P*-value may be exact (obtained from counting permutations), an approximation based on asymptotes, or derived from a model by repeated simulation. It then has to be reported and interpreted. Too many scientists — and editors —

take the line you reproach and use statistical significance as a criterion of importance.

In addition, significance is calculated in respect of a null model, chosen by the researcher and often in the knowledge that it is untenable. Why would you make measurements to compare groups if you expected to find no differences? A small *P*-value may therefore be pure fiction as a measure of knowledge gained. This comes on top of any undisclosed history of data selection and of cherry-picking results during the data analysis.

Conversely, numbers obtained from small surveys rarely demonstrate clear-cut (significant) results for individual questions, and a pattern of non-significant results in an expected direction across a range of questions could still be worth reporting as indicative. When the null hypothesis is a straw man, it may be more interesting not to be able to demonstrate the anticipated effect — for example, in a pay survey that finds no gender differences.

I endorse your view that what may seem to be sophistry is a crucial distinction. Compare, for example, the statement "The observed differences could occur 5% of the time if the true effect is zero" with the statement "The probability that the true effect is zero is 5%". Not only is the latter statement wrong, it does not match the scientific question, which should be to estimate, at a given probability, the minimum size of the effect. Another common variation is to report "no differences between groups" on the basis of *t*-tests that check for a difference only between the group means.

For scientists, talking statistics can be more dangerous than what your interviewee described as "talking Swahili in Louisiana" — unless they grasp the grammar as well as the words.

R. Allan Reese Centre for Environment,
Fisheries and Aquaculture Science,
The Nothe, Weymouth DT4 8UB, UK
e-mail: allan.reese@cefasc.co.uk

Biocultural diversity should be a priority for conservation

SIR — In his timely Opinion article on the place of botanic gardens in contemporary society ('Beyond the greenhouse' *Nature* **455**, 596–597; 2008), Mike Maunder notes that gardens are, rightly, "embracing their cultural identity". But it is not enough simply to celebrate the cultural and natural heritage of a place.

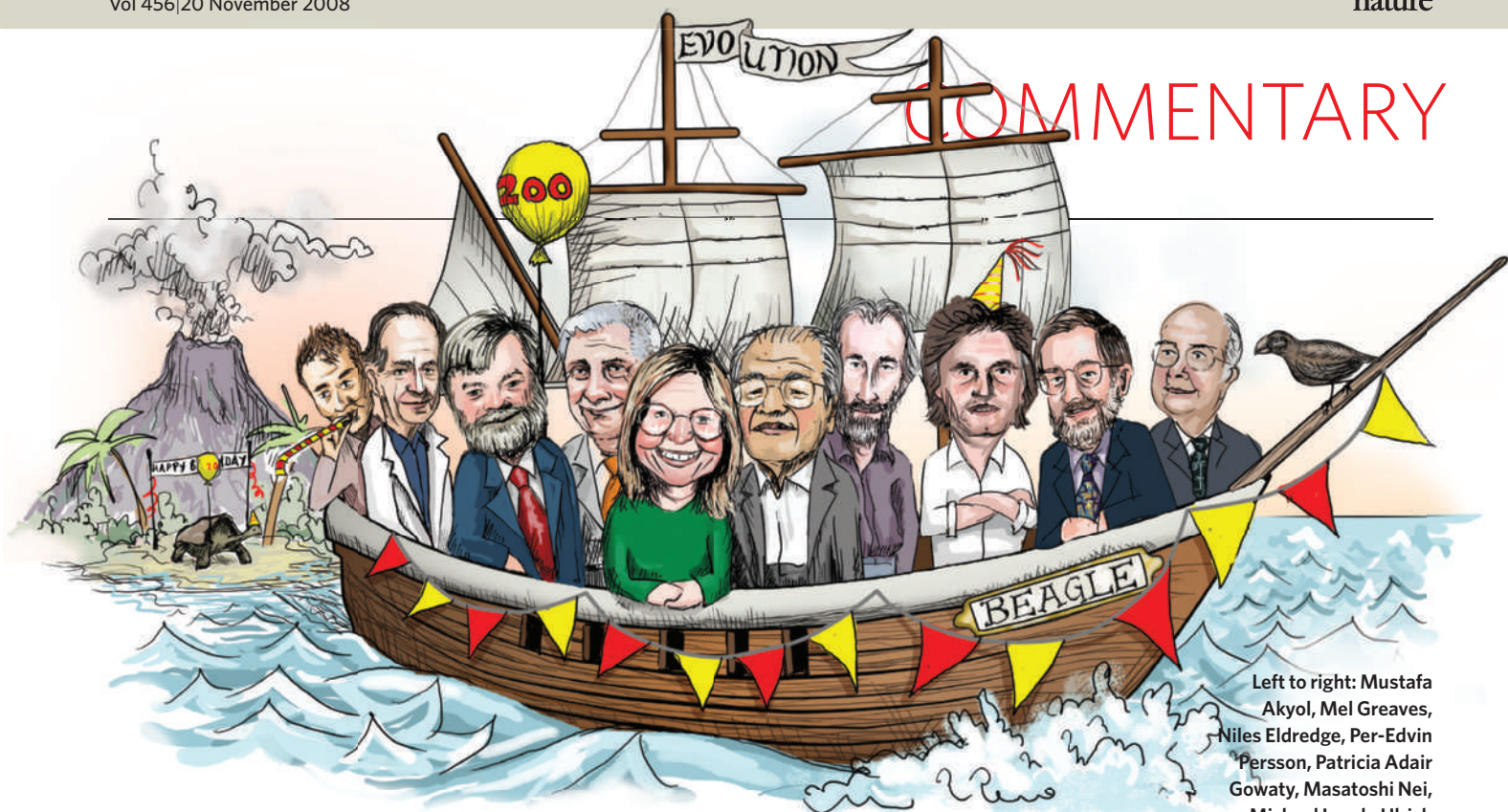
As bodies such as Terralingua and the United Nations Educational, Scientific and Cultural Organization have noted, cultures and languages are becoming as endangered as species and habitats. Superimposing maps of biodiversity hotspots and areas of cultural and linguistic diversity reveals an extensive degree of overlap, implying that similar threats are common to both (see also W. J. Sutherland *Nature* **423**, 276–279; 2003).

As biological diversity is eroded, key elements of cultural traditions, practices and language are lost. Conversely, as cultures and languages are lost, we lose irreplaceable information about the natural world, as well as notions and philosophies of place, time and humanity.

Biological conservation and cultural conservation must therefore be considered simultaneously within a 'biocultural diversity' framework. In fact, I might argue for the primacy of culture in setting conservation priorities.

Many botanical gardens are developing new education, research and collections-based programmes on plant conservation. These can become conservation leaders by explicitly incorporating cultural and linguistic conservation into their programmes and engaging local cultural practitioners.

Christopher P. Dunn Lyon Arboretum,
University of Hawai'i, 3860 Mānoa
Road, Honolulu, Hawai'i 96822, USA
e-mail: cpdunn@hawaii.edu



Left to right: Mustafa Akyol, Mel Greaves, Niles Eldredge, Per-Edvin Persson, Patricia Adair Gowaty, Masatoshi Nei, Michael Lynch, Ulrich Kutschera, Randolph Nesse, Ismail Serageldin.

Great expectations

A new path for evolution? A truce in the culture wars? Here's what a selection of readers told *Nature* they expect from Darwin 200.

Patricia Adair Gowaty

Distinguished professor, Department of Ecology & Evolutionary Biology and Institute of the Environment, University of California, Los Angeles, USA.



Darwin200

One sign of enhanced public understanding of Darwin and the nature of science, will be quicker resolution of continuously re-emerging controversies between the scientifically literate and 'creation scientists'. Other signs will include enhancements of public debate about scientific discovery, about funding for science, and policies that result from scientific discovery. There will be fewer vapid press claims about the implications of scientific studies, and an enhanced awareness of the roles of evidence and the control of bias in decision-making of all kinds.

Ismail Serageldin

Director, New Library of Alexandria, Alexandria, Egypt.

Copernicus knocked out the centrality of Earth in our view of the Universe, and Darwin knocked out the special status of humans as a species in the diversity of life on this planet. Both were vilified and attacked by bigots. Both played a central part in allowing us to

understand the reality of where we live and who we are.

Indeed, in the grand scheme of things, Darwin has given us the remarkable means to redefine the role of humans. We are the only species capable of appreciating the diversity of life and the fragility of

our ecosystems, and the only one to recognize our responsibility to change our behaviour in order to safeguard life and the world we live in.

During the Middle Ages, the Muslim world showed remarkable openness to the contrarian view and an appreciation of evidence-backed science. The Muslim world would gain much by reclaiming its legacy of open-minded pursuit of knowledge, and the Darwin celebrations could be just the catalyst to help us rediscover that tradition.

Per-Edvin Persson

Director, Heureka, the Finnish Science Centre, Vantaa, Finland.

I dream that the majority of the world's population will understand that evolution is the process by which diversity of life is maintained on this planet.

We would know this has happened by witnessing a diminished number of attacks on science, and the theory of evolution in particular, from non-scientific sources. The number

of fundamentalist believers in verbatim creation will have diminished and given way to an understanding that science and religion may coexist but that they should not be mingled. The world would accept that religion is religion and science is science and let both live in peace.

Niles Eldredge

Division of Paleontology, the American Museum of Natural History, New York, USA.

Biological phenomena that bear on evolution occur at such a mind-boggling spectrum of spatio-temporal scales that communication — hence integration — is harder now than it was in Darwin's day. Darwin himself may have been the last to have had an adequate grasp of the geology, palaeontology, zoology and botany of his day, to be able to frame something of a unifying picture. The rest of us yell at each other from increasingly widening chasms between buildings on campuses.

No one is at fault here for a lack of communication across disciplines. Do I think some of the big meetings planned for 2009 will help bridge this? One can but hope. But I doubt that Darwin's 200th birthday anniversary will manage to spur us collectively on well enough to get the job done that hasn't been successfully addressed for 150 years.

N. SPENCER

Michael Lynch

Distinguished professor, Department of Biology, Indiana University, Bloomington, USA.

My primary concern about the Year of Darwin is that the view that evolution is simply natural selection will be perpetuated further. This concern is motivated by increasingly frequent arguments, being made by people outside the field of evolutionary biology, that the entire enterprise needs overthrowing. In fact, a lot has happened in the past 150 years, and the basic theoretical framework of evolutionary biology is rock solid. There is not a single observation in cell, molecular or developmental biology that has caused a ripple in our basic understanding of evolutionary principles. This is not to say that we don't need molecular, cell or developmental biologists to complete our understanding of evolution — we need this more than ever — nor is it to say that there aren't a lot of unsolved problems.

Thus, what I would most like to see happen in the field is a true merging of the above-mentioned fields with evolutionary biology. It has long been clear that much of what we see in biology cannot be explained in terms of natural selection alone, yet we continue to witness an unwarranted proliferation of adaptive stories, in some cases extremely bizarre ones, to explain every aspect of existing and extinct biodiversity. What needs to be accomplished will take more than 12 months. More realistically, it will require the education of a new generation of scientists in the basic principles of evolutionary theory that have emerged since Darwin.

Masatoshi Nei

Director of the Institute of Molecular Evolutionary Genetics, Pennsylvania State University, University Park, USA.

We need a new evolutionary theory for the evolution of observable characteristics (phenotypes) of organisms. In this theory, mutation will play an important part, because it is mutation that generates innovative characters. The role of natural selection is merely to save beneficial mutations and eliminate deleterious ones. Because a large number of genes are involved in any phenotypic character, and evolution occurs without purpose, there must be large components of neutral or nearly neutral changes in phenotypic evolution. I doubt that this paradigm shift will be accomplished within one year, but the Year of Darwin could prompt the beginning.

Ulrich Kutschera

Professor of plant physiology and evolutionary biology, Institute of Biology, University of Kassel, Germany.

In 1859, when Darwin published his theory of descent with modification, the emerging evolutionary sciences were still mixed up with the idea of divine creation and William Paley's argument for design. Most of the nineteenth-century books dealing with the origin of life contain this pre-Darwinian biotheology. Darwin was the first to present an entirely naturalistic set of interpretations of the empirical data that provided evidence for evolution. Moreover, he discovered, independent of Alfred Russel Wallace, that natural (and sexual) selection in populations of animals and plants is a major 'driving force' for evolutionary adaptations and diversifications over thousands of subsequent generations. I hope that, by the end of 2009, Darwin's classical theories as well as his philosophical imperative — the strict separation of scientific facts from religion — will be accepted by the general public.

The clearest signs as to whether or not this message has reached the target audience of anti-evolutionists will be the acknowledgment that macroevolution

is a documented fact and not 'only a theory'.

Mustafa Akyol

Columnist for the *Turkish Daily News* and blogger for The White Path.

One thing I would really like to see is de-ideologization of Darwinism. By that, I mean the separation of Darwinism and some of the philosophies, atheism in particular, that are advanced by using this theory.

If Richard Dawkins and Daniel Dennett proclaim that their atheism is a matter of philosophical choice, not a direct outcome of 'science', and particularly Darwinian evolution, that would be a major sign that this is happening. But I am not holding my breath. I would rather expect to see more from scientists who think that evolution is compatible with their theistic faith.

A good example would be Simon Conway Morris, who thinks that evolution follows a pattern that points to a meaningful, not a meaningless, Universe. In many countries, including Turkey where I live, such views are hardly known. Neither is Darwin's comment on life on Earth as being "breathed by the Creator". If the Year of Darwin helps in discovering this often neglected side of the great naturalist, it will be a big achievement.

Randolph Nesse

Professor of psychiatry and psychology, the University of Michigan, Ann Arbor, USA.

Medicine has a blind spot for evolutionary biology. Most doctors graduate thinking that selection works mainly for the benefit of groups, that it cannot occur after menopause, that ageing results from parts inevitably wearing out and that most of the body's vulnerabilities exist simply because selection is too weak. Correcting such misconceptions requires medical curricula reform.

I am just back from a year of work with a group at the Berlin Institute of Advanced Study that considered realistic recommendations. Two are achievable in the Year of Darwin. First, national scientific organizations, such as the US Institute of Medicine, should convene groups to recommend steps that will bring evolutionary biology fully to bear on problems of human health. If they recommend that medical-school certification examinations ask questions about evolutionary biology, curricula will change quickly. Second, all schools of medicine, nursing and public health should adopt policies to ensure that their students and researchers are able to use all the tools and concepts evolutionary biology provides. We have 12 months.

Mel Greaves

Chairman, Section of Haemato-Oncology, the Institute of Cancer Research, Sutton, UK.

I would like to see both clinicians and epidemiologists recognizing that vulnerability to common diseases of affluent societies, such as diabetes, obesity, cancer and age-linked degenerative conditions, is a bequest of our evolutionary history — as mismatched with our modern life-styles. For epidemiologists, persuasion will not come from continued advocacy or polemics but from irresistible data — perhaps in the form of further genome-wide association studies implicating common allelic variants in susceptibility to common disease.

Moreover, I hope and expect to see the ongoing comprehensive, full genomic scrutiny of cancer providing detailed route maps of the evolutionary trajectories of cancer stem cells from initial emergence through to metastatic dissemination and drug resistance. The power of computational biology will have to be harnessed to manage the rich and dynamic complexity that will emerge. ■

See more responses and join the discussion at <http://tinyurl.com/57h7js>. See also pages 281, 295 and www.nature.com/darwin.

BOOKS & ARTS

Colonies that conquer

A landmark book documents how social insects form highly successful colonies that mimic a single organism, explains **Manfred Milinski**.

The Superorganism: The Beauty, Elegance, and Strangeness of Insect Societies

by Bert Hölldobler and E. O. Wilson

W. W. Norton: 2008. 576 pp.

\$55.00, £30.00

Being just one cell, a bacterium has to be a jack-of-all-trades. By contrast, multicellular organisms, such as humans, profit from a division of labour between their specialized organs. The digestive, immune, circulatory and other systems all cooperate to help the gonads to maximize the reproductive success of that organism in competition with other organisms. Would a village of human inhabitants, in competition with other villages, profit similarly if its foragers, defenders and communicators volunteered to help the reproducers maximize their reproductive success? Certainly not. But many social insects do just this, as Bert Hölldobler and Edward Wilson beautifully portray in their new masterpiece *The Superorganism*.

A scientist from another planet would describe the most widespread, highly social form on Earth as having six legs, two antennae and three body parts, living in groups of up to 20 million members. Ants, bees and termites have evolved into superorganisms — colonies with a strict division of labour, in which one or a few individuals reproduce while many

nonbreeding workers with various specializations carry out tasks altruistically.

Eighteen years after their Pulitzer prize-winning book *The Ants*, the authors have put together a wealth of new research to present the rich and diverse natural history facts that define a superorganism's traits — the colony as a self-organized entity and a subject of natural selection.

Most of us are familiar with the colony life of honeybees. One female, the queen, is the mother of all offspring, including her 'busy as a bee' worker daughters who build, clean and guard the nest, feeding the larvae to raise many sisters and brothers in competition with other colonies. Worker bees all look alike, each practising all professions sequentially, except reproduction.

Many ants, however, have castes of workers with different morphologies; one example is the worldwide ant genus *Pheidole*, with its unusually large-headed soldiers and many minor workers.

Cooperation in colonies can be very sophisticated. The aerial pavilions of the weaver ant *Oecophylla* are constructed of leaves bound together by silk. One group of workers cooperates to pull the edges of the leaves together. Another group brings out large, larval sisters and wags them bodily back and forth across the newly created seams. With each sweep, the larvae emit sticky threads of silk over the seams to



Darwin200



hold the leaves in place. Hölldobler and Wilson present numerous further fantastic examples.

The evolution of this division of labour is a fascinating topic. All colony members have the same genotype for caste formation, but in response to environmental conditions, these genes turn growth on and off specifically towards one caste or another. That which is needed is produced, dictated by colony economy.

Why are insect workers so altruistic that they give up their own reproduction to help others to reproduce? Natural selection favours traits that help their possessor to have more offspring. Thus, organisms that help others at the expense of their own offspring are difficult to explain. Why should the brain cells of an organism help her gonads to produce offspring? Simply, because the genes of brain cells and gonad cells are identical — there is no conflict of interest among cells of an organism about which cells should reproduce. Any division of labour that boosts reproductive success should evolve. Human inhabitants of a village, however, differ in their genes and would not agree on who should reproduce and who should raise the reproducers' offspring.

Colonies of social insects are similar to the cells of an organism; they start as a family. Being altruistic has low costs because helping the mother to produce more siblings is genetically equivalent to producing one's own offspring. The former strategy can be more efficient through division of labour. The basic



Royal servants: the queen bee (centre) produces offspring while her worker daughters do the rest.



D. CLYNE/PHOTOLIBRARY.COM

Seamless cooperation between weaver-ant workers allows them to build elegant aerial nests using leaves.

idea that altruism towards kin should evolve goes back to the founding fathers of population genetics, Ronald Fisher and J. B. S. Haldane. They proposed that altruism is more likely to evolve the more the helper and helped are related, and the more efficient that help is. William D. Hamilton's seminal work on inclusive fitness theory made the idea explicit.

Most superorganisms are found among the Hymenoptera, or bees and ants. Owing to their special sex-determining mechanism, females develop from fertilized eggs and males from unfertilized eggs. Because their father has only one set of chromosomes, the sisters are 75% related to each other instead of the usual 50%. Sisters are less related both to their brothers (25%) and to their own offspring (50%), if they had any. This explains why, in all hymenopteran colonies, workers are the daughters and never the sons of the queen. This high relatedness among sisters is, as Hölldobler and Wilson point out, not a necessary but a helpful prerequisite for crossing the so-called eusociality threshold, when a worker propagates more of her own genes through related sisters by staying in the colony and helping than by trying on her own. Termites and even one mammal, the naked mole-rat, have become superorganisms despite their normal mode of reproduction. Maximizing efficiency through division of labour is of key importance besides relatedness. Thus, organized

groups beat solitary individuals in competition for resources, and large groups beat smaller ones of the same species. After the point of no return — when an anatomically distinct worker caste exists — a lower level of within-colony relatedness can be tolerated, as long as this is higher than the level of relatedness between colonies.

Hölldobler and Wilson present a thorough account of the selective forces that led to the evolutionary construction of superorganisms. They favour 'multilevel selection', a combination of within-colony selection on individuals helping their relatives and between-colony selection, so-called group selection. Colonies compete and the best adapted win. The logic of group selection — groups with altruists are more productive than those without — is compelling. However, the conditions that allow altruists to survive within large groups are still annoyingly demanding. The trick of superorganisms may be that altruists benefit anyway through efficient sister production. Recently, a fierce debate has started among theorists, questioning whether inclusive fitness models can easily incorporate all mathematical intricacies of multilevel selection. Whatever the solution might be, superorganisms certainly have it.

Communication is highly evolved in superorganisms. Besides the celebrated dance language of honeybees, ants have become the

insect geniuses of chemical communication with numerous pheromones. This and many other fascinating topics — such as the combined phylogenetic and ecological explanation for the immense diversity of ants, or the beauty and functionality of ant nest architecture — are presented in depth, but always in a lucid, entertaining language combined with magnificent line drawings and photographs. My favourite chapter highlights the leafcutter ants, "Earth's ultimate superorganisms". As well as having the most complex communication system known in animals, they have the most elaborate caste systems, build air-conditioned nests, and made the transition from a hunter-gatherer existence to agriculture — by growing their fungal food source in their nests using elaborate techniques — long before humans did.

This monograph is a state-of-the-art review of the organization and evolution of social insect societies. Detailed references on almost every page help to track information back to its roots. This book will fascinate everybody interested in biology, social sciences, economy or simply natural history. ■

Manfred Milinski is director of the Max Planck Institute for Evolutionary Biology, Plön, Germany. e-mail: milinski@evolbio.mpg.de

See News Feature, page 296.



HMS Beagle off Tierra del Fuego, by the ship's artist Conrad Martens.

NATIONAL MARITIME MUSEUM, LONDON

Darwin: heading to a town near you

The theory of evolution challenges artists and philosophers as much as scientists. **Joanne Baker** rounds up the many forthcoming events worldwide that examine Darwin's life, his work and reactions to it.

Charles Darwin's ideas speak to the essence of what it is to be human. Different cultures have absorbed and expressed this in different ways. His impact on science and culture is reflected in the array of exhibitions, events and books planned for 2009, in celebration of the bicentenary of his birth and the 150th anniversary of the publication of his seminal work, *On the Origin of Species*.

Journeys

A profound experience in Darwin's life was his five-year voyage across the southern oceans aboard the survey ship HMS *Beagle*. In 2009, a lucky few will be given the chance to retrace Darwin's steps. A modernized replica of the ship will carry a crew of scientists, teachers and students along the route of the 1831–36 voyage (www.thebeagleproject.com). Their experiments will be beamed to classrooms worldwide. Watching from space, astronauts aboard the International Space Station will help the crew to investigate the biology of plankton blooms, coral reefs and other ocean-surface and terrestrial ecosystems.

If you have less time and US\$60,000 to spare, then you could follow Darwin's route in three weeks by private jet. Stanford University is offering this bespoke trip accompanied by three science professors, a physician and a chef. Departing from California in January, the Boeing 757 and its 92 passengers will visit the Galapagos and Easter Islands in the Pacific, Tierra

Del Fuego in South America, Tasmania, Mauritius, Cape Town in South Africa and the Cape Verde Islands off Africa's west coast. Carbon offsetting does not seem to be included.

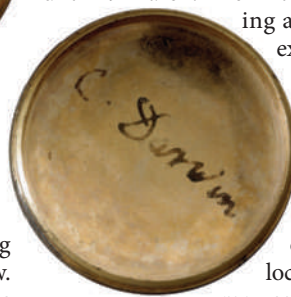
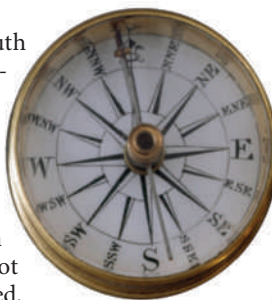
Exhibitions

Artefacts from Darwin's original voyage feature in many exhibitions, including *Darwin* at London's Natural History Museum, which runs until 19 April. In a week-long festival in Cambridge, UK (5–10 July; www.darwin2009.cam.ac.uk), that city's museums will display Galapagos finches (as pictured, right), barnacle microscope slides, fish, beetles, rocks and scientific instruments. Papers by Darwin and other evolutionary thinkers will be on display at the British Library, London, in *Darwin and the Story of Evolution*, from 10 December 2008 until 22 March.

Beyond Darwin's native United Kingdom, where most attention is understandably focused, city-wide festivals are also popular — including those in San Francisco, Philadelphia and Houston in the United States. Museums are collaborating to spread Darwin material globally. The American Natural History Museum's successful *Darwin* exhibition — which first went on show

in New York in 2005 and has travelled across the United States, South America and Japan — forms a template for some museum exhibitions, including those at London's Natural History Museum and The National Museum of Australia in Canberra, from 10 December 2008 until 29 March. The British Council is sharing a modular Darwin exhibition with more than 25 countries, including South Korea and Hong Kong, China.

In nations where he spent less time or where the ideas of local influential figures may compete, exhibitions focus more on the science than the man. *Darwin's Evolution* at the Calouste Gulbenkian Foundation in Lisbon, Portugal, runs from 12 February until 24 May and will examine evolutionary theory from its origins to current genetics research. Other major events are planned at the Museum of Natural History (*Naturkundemuseum*) in Berlin, at the Toulouse Museum in France and at the State Darwin Museum in Moscow, which was founded in 1907 to celebrate the role of Russian scientists in developing evolutionary theories. The wider social effects of Darwin's ideas will be explored at an international conference in November 2009 at the Bibliotheca Alexandrina in Alexandria, Egypt.



ENGLISH HERITAGE PHOTO LIBRARY

NATURAL HISTORY MUSEUM, LONDON

Imagery

Artists' responses to Darwin's theories of human origins will be displayed at the Schirn Kunsthalle in Frankfurt, Germany. *Darwin: Art and the Search for Origins* (5 February until 3 May) will highlight the struggle for survival portrayed in paintings such as *Passion Flowers and Hummingbirds* by Martin Johnson Heade, depictions of the emergence of life by Arnold Böcklin and Odilon Redon, and Max Ernst's artistic explorations of 'deep time'. *Endless Forms: Charles Darwin, Natural Science and the Visual Arts* shows at the Yale Center for British Art in New Haven, Connecticut (12 February until 3 May), and then at the Fitzwilliam Museum, Cambridge, UK (16 June until 4 October). It will explore the importance of visual imagery in the development of Darwin's ideas and the effect of his theories on artists such as J. M. W. Turner, Edgar Degas, Claude Monet and Paul Cézanne.

Performances

Darwin's quest to understand human origins also motivates composers, musicians and playwrights. New music includes a cycle of folk songs and the classical *Age of Wonders* by British composer Michael Stimpson, which will premiere in London on 21 January before touring to Australia in September. Performances of *Re:Design*, a play dramatizing the correspondence between Darwin and the Harvard University botanist Asa Gray, will include a show in Cambridge, UK, on 10 July. On film, actors Paul Bettany and Jennifer Connelly will play Darwin and his wife Emma in the feature *Creation*, directed by Jon Amiel.

Darwin's ideas were controversial in his time and remain so today. It is worth attending some of the many anniversary events to explore how his thinking has been interpreted, and what it says about how we see ourselves. ■

Joanne Baker is *Nature's* Books & Arts editor.

For more on Darwin, see also pages 281 and 295, and online at www.nature.com/darwin.

For Darwin events near you, see www.darwinday.org and www.darwin200.org.

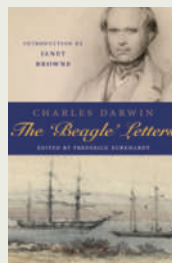


A Down House bookshelf

ON THE SEAS

Darwin's five-year voyage on HMS *Beagle* was spent documenting nature across the Southern Hemisphere. His full notes are now compiled in *Charles Darwin's Notebooks From the Voyage of the Beagle*

by historian John van Wyhe with Gordon Chancellor and Kees Rookmaaker (Cambridge Univ. Press). In *Charles Darwin: The Beagle Letters*, edited by Frederick Burkhardt (Cambridge Univ. Press), Darwin's correspondence charts his changing ideas and everyday seafaring experiences. Alan Gibbons offers young readers a cabin boy's perspective in *Charles Darwin* (Kingfisher). And in *The Voyage of the Beagle: Darwin's Extraordinary Adventure Aboard Fitzroy's Famous Survey Ship* (US Naval Institute Press), James Taylor collates the ship's plans, biographies of Darwin and Captain Robert Fitzroy, photographs, artefacts and journal extracts from the voyage.

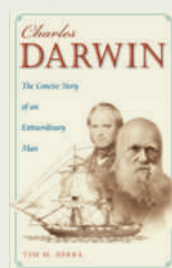


ON EVOLUTION

In *Charles Darwin: The Concise Story of an Extraordinary Man* (Johns Hopkins Univ. Press), Tim Berra

describes Darwin's revolutionary scientific work, its effect on modern biology and the influence of evolutionary theory on Western thought. *The Rough Guide to Evolution* by Mark Pallen (Rough Guides) provides a popular summary, including the cultural impact of Darwin on music, plays and novels, and a list of Darwin bicentenary events.

Charles Darwin's Shorter Publications 1829-1883 by John van Wyhe (Cambridge Univ. Press) includes Darwin's theory of evolution by natural selection and more than 70 newly discovered items. *The Cambridge Companion to the 'Origin of Species'*, edited by Michael Ruse and Robert J. Richards (Cambridge Univ. Press), draws on expert opinion to provide the religious, social, literary and philosophical contexts in which the *Origin* was composed.



ON THE MAN

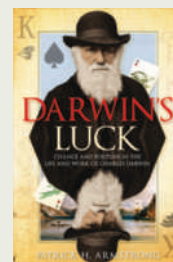
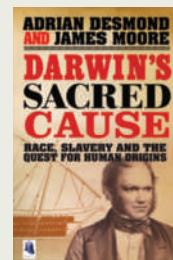
In their controversial forthcoming biography, *Darwin's Sacred Cause* (Allen Lane), Adrian Desmond and James Moore will argue that Darwin's hatred of slavery fuelled his search for a theory of human origins. After examining his manuscripts and letters, they conclude that this disgust empowered the conservative man to come up with theories that were seen as radical by his contemporaries.

Darwin's morality was also influenced by his wife's religious beliefs, according to Deborah Heiligman in her biography aimed at younger readers, *Charles and Emma: The Darwins' Leap of Faith* (Henry Holt). His own poor health may also have coloured his thoughts, says Ralph Colp in *Darwin's Illness* (Univ. Press Florida).

Or he was plain lucky, according to Patrick H. Armstrong. *Darwin's Luck: Chance and Fortune in the Life and Work of Charles Darwin* (Hambledon Continuum) asks to what extent Darwin took the wrong scientific paths, even if he eventually came to the right conclusions.

Darwin's exploration closer to home is the subject of *Darwin's Island* (Little, Brown). Biologist Steve Jones visits the British landscapes that influenced 'the sage of Kent' as much as the Galapagos Islands, and brings his work up to date. Michael Boulter's book, *Darwin's Garden* (Constable & Robinson), describes Darwin's relationship with his own garden and the experiments he carried out in his greenhouse, explains Beverley Glover in her review (see *Nature* 454, 944-945; 2008).

Darwin shares his birthday celebrations with Abraham Lincoln. In *Angels and Ages: A Short Book About Darwin, Lincoln and Modern Life* (Random House), Adam Gopnik sees both as thinkers who made and experienced great changes in society. They saw the shift away from faith and fear to the embrace of reason, argument and observation not merely as intellectual ideals, but as a way of life. ■



ESSAY

Birthdays to remember

Anniversaries of Charles Darwin's life and work have been used to rewrite and re-energize his theory of natural selection. **Janet Browne** tracks a century of Darwinian celebrations.

Anniversaries are big business in the cultural world and have long been convenient events for promoting agendas. Tourism, commerce, education; all these can be boosted in the name of an anniversary.

In science, anniversaries help us to explore the implications of truly important achievements, confirm shared ideas, highlight the value of key players and look forward to new problems to resolve. As we fast approach 2009, the bicentenary of Darwin's birth and the 150th year since the publication of *On the Origin of Species*, it is worth remembering that scientific anniversaries also provide an opportunity to push an agenda, and even to adapt the past, so telling us what we like best to hear.

It is becoming clear to historians that Darwin commemorations held at various points over the past century have been used to re-frame and re-establish the relevance of natural selection at delicate moments. Interest in Darwinism rose and declined noticeably after his death. The idea of natural selection fell out of favour around 1900, when the field of genetics came to life, and again in the 1950s when molecular biology emerged. Each of these new themes probably made existing biological explanations look old-fashioned. An increasing amount of evidence suggests that Darwinism needed intellectual reboots at strategic times in the twentieth century to become the theory that it is today. The urgency of these former scientific rallying points has been obscured by the success and high profile of modern selection theory.

Funeral as propaganda

The first such posthumous rallying event was Darwin's funeral in April 1882. This was an occasion for national glorification. As biographer James Moore described in a 1982 article¹, active scientific reformers in the Royal Society of London — including Thomas Henry Huxley, Francis Galton and John Lubbock — moved rapidly after Darwin's death to obtain permission to bury him in Westminster Abbey, one of the most prominent religious buildings in Great Britain. This was a far cry from the simple funeral that Darwin had expected in his village church. Eight days later his coffin was carried with great ceremony through the nave of the abbey. The funeral service and many



Darwin200

obituaries stressed that Darwin was not an atheist. He was instead described as a good man, committed to truth and honesty. This was true, but it was also valuable propaganda at a time when relations between science and religion were intensely fraught. The men of the

Royal Society used Darwin's funeral as a way to reassure their contemporaries that science was not a threat to moral values, but rather was becoming increasingly important in the modern world.

In 1909, 100 years after Darwin's birth, a number of celebrations were held in Europe and North America. One of the more memorable took place at the University of Cambridge, UK, where two of Darwin's sons were professors. As documented by historian Marsha Richmond of Wayne State University in Detroit, Michigan², visiting scientists were given honorary degrees, Darwin's former college rooms in Christ's College were opened, manuscripts were displayed and a substantial volume of essays by noted biologists was published.

Less widely appreciated, however, is that this meeting took place at a time when Darwinism as a biological explanation was going through a bad patch. Some were beginning to think that the notion of gradual change simply wasn't needed. After the word 'genetics' was coined in 1905, much work focused on the mutation of chromosomes. It was assumed by great experimentalists such as Thomas Hunt Morgan and Theodosius Dobzhansky that the breaking and rejoining of chromosomes, or some other type of chromosome alteration, was sufficient to generate permanent change in an organism. Thus, new forms could emerge *de novo*, without selective pressure and adaptive success.

Darwinism had a troubled place in palaeontology, too. The subject was then leaping forwards with remarkable fossil discoveries, and it was increasingly possible to propose ancestral trees with some certainty. Yet it seemed as though fossil animals had progressed along particular lines of development, as if with some final purpose in mind — a notion that put teleology back into biology only 50 years after Darwin had done so much to try to remove it.

At the same time, conventional evolutionary

biology seemed to be losing any sense of unity, potentially diluting the power of Darwin's all-embracing idea. Biometricians such as Karl Pearson focused on a statistical view of populations to study evolution; pioneering ecological thinkers such as Eugen Warming saw the key issue as organisms' struggle directly against the environment.

Darwinism as set out by Charles Darwin seemed increasingly sidelined. The 1909 commemorations, organized by a small group of naturalists and Darwin family members from the University of Cambridge, provided a way to reassert the primacy of natural selection against other evolutionary rivals.

American stage

Then, in 1959, a hundred years after *On the Origin of Species* was published, it was the turn of the United States and the new 'biological systematics', argues historian Betty Smocovitis of the University of Florida in Gainesville³.

This Darwin anniversary was held at the University of Chicago in Illinois, in a symposium that pointedly celebrated the integration of genetics and population statistics with selection theory. Ten years earlier, this integration had almost taken the form of a political treaty. Putting it bluntly, field naturalists were eager to

re-establish their value in an increasingly laboratory-based world. Prominent naturalists such as Ernst Mayr managed to get geneticists and statisticians to agree that evolution could take place on three levels: in molecules; in the

flow of genes through populations; and in the environmental world of organisms undergoing competition and natural selection. In 1942, Julian Huxley invented the phrase 'modern synthesis' to combine genetics with natural selection, and Mayr's key work within this synthesis, *Systematics and the Origin of Species from the Viewpoint of a Zoologist* (Columbia Univ. Press), was published.

The delegates at Chicago did more than celebrate a new union of the biological sciences. They in effect created modern Darwinism by emphatically rejecting any form of Lamarckism — the doctrine that an organism could pass on traits acquired through environmental conditions during its lifetime. In 1959, socialist Russia had only recently withdrawn

"Darwin would surely be amazed by how differently we have chosen to celebrate his anniversaries."

DARWIN ARCHIVE, CAMBRIDGE UNIV. LIBRARY



Party snaps: Darwin family members gathered at the 1909 celebrations in Cambridge, UK (above); Darwin's grandson, the physicist Charles Galton Darwin (below left) on a 1959 televised US talk show with colleague Julian Huxley (centre) and politician Adlai Stevenson (right).

L. BALTERMAN/TIME & LIFE PICTURES/GETTY IMAGES



from Lamarckism in genetics, and the idea was strongly associated in US minds with the cold-war struggle. The delegates also rejected the idea that the fossil record shows signs of directed evolution, and expanded Darwinian thought to cover the evolution of mind and behaviour. During the conference, Julian Huxley, the grandson of Thomas Henry Huxley, gave a secular sermon in the style of his grandfather, and provocatively declared that religious belief was merely a biological feature of evolving mankind.

Evolution of a theory

The story of Darwin's finches as a classic example of evolution in action also became more widely known at this point. Biologist David Lack had been encouraged by Mayr and Julian Huxley to spend time studying the finches on the Galapagos archipelago, first with a view to their ecology, but then as an example of beak diversification according to the availability of foodstuffs and lack of other bird

competition. It was only after this, according to Frank Sulloway of the University of California, Berkeley, that the finches sketched by Darwin became collectively known as Darwin's finches, and were held up as the first and most remarkable evidence of evolution in real organisms in a natural setting.

In fact, much of what we know about Darwin and Darwinism, including his celebrity status, is the result of the 1959 celebration in Chicago. Plans were announced for a Darwin memorial park on the Galapagos Islands, which tied in with international pressure on Ecuador to restrict commercial fisheries and allowed the islands to become a designated World Heritage area. Darwin's original books and papers were preserved through munificent bequests from US foundations and the Darwin family, and several noteworthy biographies were published.

In 2009 we shall see an unprecedented number of Darwin commemorations, including scientific symposia, books (see page 323), postage stamps and a big-budget film (see page 322).

Will these forthcoming activities have a veiled agenda, as did those of the past?

Without question, biologists will pay tribute to the theory itself — a magnificent achievement that lives on as the central organizing concept of modern biology and much else besides. In 1859, the clarity and insight of Darwin's vision in *On the Origin of Species* provided real answers to the biggest natural-history problems of his day. Since then, the idea of evolution by natural selection has been restructured from the bottom up, to account for new fields such as genetics. In this way it has persisted, and shown remarkable explanatory power for nearly 150 years.

But biologists will also surely use the occasion, once again, to affirm the truth and elegance of Darwinism in the face of criticism, this time from those who prefer a creationist view of the world. Evolution by natural selection has suddenly become a highly contentious idea, especially in the United States. Creationist proponents abound in the US school-board system, opinion polls highlight the public's belief in a divine origin for humankind, and ideas about intelligent design are widely circulated. Against this, Darwin has become the figurehead for rational, secular science, and Darwinism the main target of the fundamentalist movement spreading across the globe. Attacks extend beyond arguments over the Bible. To criticize Darwinism is a forceful way to express anxieties about the growing power of modern science and the perceived decline of moral values in society. To try to poke holes in Darwin's argument is to express dislike not just for evolutionary theory but also for science itself.

There is some irony in this situation. Looking back to Darwin's funeral in 1882, Darwin's Christian qualities, his stature as a man of truth and honesty, were brought to the fore. He was celebrated as a man whose religious doubts were an integral part of his wisdom and insight; few critics made personal attacks on his social virtues. Now, his heroism in modern science is seen by many as an offence to religious values. It goes to show just how diversely Darwin and his theory have been perceived and used over the years. Darwin himself would surely be amazed by how differently we have chosen to celebrate his anniversaries. ■

Janet Browne is a historian of the life sciences at Harvard University, and the author of books including *Darwin's Origin of Species: A Biography*. e-mail: jbrowne@fas.harvard.edu

1. Moore, J. R. *Biol. J. Linn. Soc. Lond.* **17**, 97–113 (1982).
2. Richmond, M. L. *Isis* **97**, 447–484 (2006).
3. Smocovitis, V. B. *Osiris Second Series* **14**, 1–66 (1999).

See also pages 281, 295, 322 and 323, and online at www.nature.com/darwin.

ESSAY



Conflict: Altruism's midwife

Generosity and solidarity towards one's own may have emerged only in combination with hostility towards outsiders, says **Samuel Bowles**.

Groups of fire ants, chimpanzees, meerkats and other animals engage in lethal conflicts. But we humans are especially good at it, killing 'outsiders' on a scale that altered the course of our evolution. Prehistoric burials of large numbers of men and women with smashed skulls, broken forearms and stone points embedded in their bones — as well as ethnographic studies of recent hunters and gatherers — strongly suggest that warfare was a leading cause of death in many ancestral populations. Yet at the same time, humans are unusually cooperative, collaborating with non-kin, for example in hunting and sharing food, on a scale unknown in other animals.

Paradoxically, the grisly evidence of our warlike past may help explain our distinctly cooperative nature.

This 'distasteful' idea is based on the evolution of what my co-authors and I have termed 'parochial altruism'. Altruism is conferring benefits on others at a cost to oneself; parochialism is favouring ethnic, racial or other insiders over outsiders. Both are commonly observed human behaviours that are well documented in experiments. For example, people from the Wolimbka and nearby Ngenika groups, in the Western Highlands of Papua New Guinea, have no recent history of violence. Yet when asked to divide a pot of money between themselves and another, they give more and keep less for themselves if the other is a member of their own group rather than an outsider.

But parochial altruism is puzzling from an evolutionary perspective because both altruism and parochialism reduce fitness or material well-being compared with what a person would gain were he or she to eschew these behaviours. Altruistic acts, by definition, confer advantages on others at a cost to the altruist. The impediments to the evolution of parochialism are more complicated, but could also be prohibitive. Hostility towards

outsiders limits an individual's choice of partners for long-distance trade, political coalitions and help during times of adversity. Like the altruist, the parochialist bears a handicap in the evolutionary race.

The solution to the puzzle may be that parochialism and altruism act synergistically. Among ancestral humans, parochial altruists may have provoked conflicts between groups over scarce natural and reproductive resources, and at the same time contributed to a group's success in these conflicts. Altruism would have facilitated the coordination of raiding and ambushing on a scale known in few other animals, while parochialism fuelled the antipathy towards outsiders. Additionally, with the development of projectile weapons, humans became adept at killing from a distance, which would have reduced the costs of aggression.

As winning groups gained territory, an increase in reproductive opportunities and political and cultural influence could have overcome the selective disadvantages of parochialism and altruism when occurring separately. When winners and losers differed significantly in their genes or cultural practices, the effects of this kind of conflict on evolution could have been substantial.

Computer wars

Support for this idea comes from artificial histories of early human evolution that my co-authors and I simulated by computer. In these simulations, we allowed groups of agents, tolerant or parochial, altruistic or selfish, to interact over thousands of generations under conditions likely to have been experienced by our Late Pleistocene and early Holocene ancestors. We designed the simulations so that violent conflict between two groups is likely if

at least one group contains a preponderance of parochialists. We also made each group's fighters the parochial altruists (non-altruists are happy to let someone else do the fighting; tolerant members prefer to stay on friendly terms with outsiders). Thus, the groups with the most parochial altruists tend to win conflicts. Our objective was to see how the frequency of warfare, and the fraction of the different types of agent, would evolve.

In millions of simulated evolutionary histories, the populations emerging after thousands of generations of selection tend to be either tolerant and selfish, with little warfare, or parochial and altruistic with frequent and lethal encounters with other groups.

Occasional transitions occur between the selfish peaceful states and the warring altruistic states. But neither altruism nor parochialism ever proliferate

singly; they share a common fate, with war the elixir of their success.

Climatic and archaeological evidence indicate that competition between groups, which underpins the process by which parochial altruism evolves, was rife during the Late Pleistocene, about 126,000 to 10,000 years ago. The extraordinary climate instability recorded in Greenland ice cores would have heightened competition for resources, forcing long-distance migrations and frequent encounters among hard-pressed and, in the later part of the period, well-armed groups. Archaeological finds, such as projectile wounds found in the skeletons on the Channel Islands off southern California, suggest that during seven millennia of prehistory, conflict among groups was especially intense during periods of environmental stress. Such conflict seems to have accounted for a much larger fraction of deaths

"The inspiring public spiritedness, courage and generosity that are distinctive of humans bear the birthmarks of a history of conflict."



G. BECKER

than occurred during Europe's just-concluded 'century of total war'.

Charles Darwin anticipated the contribution of warfare to the evolution of altruism, predicting that a tribe possessing a greater number of members ready to warn each other of danger, and aid and defend each other, would spread and be victorious over other tribes. Thus "the social and moral qualities would tend slowly to ... be diffused throughout the world". But he omitted to mention that among these 'moral qualities' would be hostility towards outsiders.

From genes to culture

A similar evolutionary logic extends to the present, working on timescales more appropriate for cultural than genetic evolution. Indeed, the modern European state was forged in the heat of warfare among the some 500 city states, bishoprics, principalities and other sovereign bodies that governed Europe half a millennium ago. Parochial conflict was the midwife of the novel institutions — tax compliance, respect for the rights of property, the rule of law — that spelled survival in the five-century-long shakeout that, on the eve of the First World War, had left just 27 states standing.

The making of Europe as we know it thus paradoxically owes something to the exploits of "animals possessing the virtues of courage and fighting, but nothing else" in the words of a twelfth-century Islamic soldier-scholar who lost his home and family to the Crusaders.

However, I do not want to oversell this 'red in tooth and claw' side of human origins. Even in periods and places where warfare was uncommon, environmental crises could have eliminated groups that failed to work together while cooperative groups survived.

Moreover, like the emergence of multicellular organisms, much of human distinctiveness got an evolutionary boost from practices that kept the lid on conflict among group members. Where this occurred, members would have tended to share common levels of

reproductive and cultural success. As a result, the evolutionary effect of competition between individuals would have receded in importance compared with that of competition between groups, giving cooperators the edge.

Practices that suppress competition within groups are common among recent hunters and gatherers. These include sharing food and information, and mating systems that prevent dominant males from monopolizing reproductive opportunities. The foraging Aché people in Paraguay, for example, shared foods such as meat and honey equally among group members, irrespective of who originally acquired them. Among our foraging ancestors, similar practices almost certainly protected altruistic and cooperative individuals from exploitation by aggressive self-aggrandizers.

Our computer simulations of evolutionary histories show that in groups adopting these so-called levelling practices, the tendency of altruistic members to be eliminated by natural or cultural selection is attenuated. Indeed, without the suppression of conflict within groups, empirically plausible levels of conflict between groups do not promote the spread of altruism.

Thus, in ancestral humans, evolutionary pressures favoured cooperative institutions among group members as well as conflict with other groups. These were complemented by individual dispositions of solidarity and generosity towards one's own, and suspicion and hostility towards others. This potent combination of group and individual attributes is as characteristic of the contemporary welfare state in a system of heavily armed and competing nations — in short, modern nationalism — as it was among our ancestors.

Legacy, not fate

The inspiring public spiritedness, courage on behalf of others and generosity that are distinctive of humans thus bear the birthmarks of a history of conflict. But our ancestors made peace as well as war. They benefited from

an interchange of goods, information, words, songs and marriage partners — anticipating the networks of risk sharing, friendship and exchange that would flourish among modern foragers.

The !Kung foragers of Namibia and Botswana, for example, exchanged goods (many of purely symbolic value) with members of other groups, often more than 100 kilometres away. This was to ensure that they and their families would be welcomed should personal conflicts or a dry water hole necessitate moving. If the !Kung could devise cultural practices ensuring the welcome of needy people from distant groups, a similar spirit of solidarity could surely exist among today's far-flung peoples just an e-mail or overnight flight away.

Moreover, humans are uniquely receptive to socialization and learning, with tolerant altruistic behaviour a frequent result. Ethnic hostility can be redirected, attenuated and even eliminated in a matter of decades or years. This has been shown by the intellectual, political and even military collaboration of Muslims, Christians and Jews that occurred in parts of Islamic Spain a millennium ago, even as the First Crusade pitted Christian against Muslim in the eastern Mediterranean; the widespread support in many countries for aid to the people of poor nations; and the recent transformation of racial attitudes among US voters. Those who joined Martin Luther King in 'We Shall Overcome' were not only buoying their spirits; they were also making a reasonable assessment of human possibilities.

Thus, even if I am right that a parochial form of altruism is part of the human legacy, it need not be our fate.

Samuel Bowles directs the Behavioral Sciences Program at the Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, New Mexico 87501, USA, and is professor in the Department of Political Economy at the University of Siena, Italy. He is the author of *Microeconomics: Behavior, Institutions and Evolution*.
e-mail: bowles@santafe.edu

See <http://tinyurl.com/5nuy3l> for further reading.
For more on Being Human, see www.nature.com/nature/focus/beinghuman.

NEWS & VIEWS

ASTROPHYSICS

A message from the dark side

Yousaf M. Butt

Both astrophysicists and particle physicists are in on the hunt for the elusive dark matter that is thought to pervade the Universe. A high-altitude balloon-borne experiment offers the latest hints as to what it could be.

Humiliating as it may sound, you, me and everything we see — the Earth, Moon, Sun and stars — may be little more than cosmic contamination. Most of the ‘stuff’ in the Universe is thought to be in the form of invisible and elusive particles of dark matter. To date, the existence of this cosmic exotica has been inferred through its gravitational effects. But on page 362 of this issue, Chang and collaborators report¹ on a surprising bump detected in the spectrum of celestial electrons that could be a more direct signal of this mysterious substance — or there may be other intriguing explanations.

For 75 years, astronomers have collected data that point to the existence of a type of non-luminous matter that outweighs normal (‘baryonic’) matter by a factor of about six. Several independent lines of evidence seem to make its reality compelling. For one, the measured rotation speeds of stars and gas within a typical galaxy are such that the galaxy would simply fly apart were it not for the gravitational anchor of copious non-baryonic dark matter. Similarly, the orbital speed of individual galaxies in a galaxy cluster also supports its presence. The slight gravitational warping of space, as evidenced by gravitational lensing (which bends the light from distant objects as the light travels towards Earth), is yet another indicator of its existence². Dark matter also seems to be a necessary ingredient for making sense of the fluctuations detected in the structure of the cosmic microwave background, the pervasive ‘echo’ associated with the young, hot Universe³.

However, alternatives to dark matter, such as a modification to the law of gravity (known as MOND⁴, for ‘modified Newtonian dynamics’), have been proposed and can also explain some of the data well. So even though there is not universal consensus regarding the reality of dark matter^{5,6}, on balance the case for its existence seems robust.

But what could these mystery dark-matter particles be? All known fundamental particles — even the ever-elusive neutrinos — are excluded as dark-matter candidates, so we are forced to look beyond the ‘standard model’ of particle physics. The most promising candidates



The ATIC balloon-borne electron detector ready for launch on the Ross Ice Shelf of Antarctica.

T.G. GUZIK

so far are WIMPs, short for weakly interacting massive particles. And the top suspects within this generic class come in two flavours: the neutralinos, which arise in supersymmetry theories, and Kaluza–Klein (KK) particles, which emerge in theories involving extra dimensions⁷ and which are named for their first proponents in the 1920s.

Unlike astrophysicists, who look up at the sky and measure the cumulative gravitational effects of dark-matter particles on enormous, galactic scales, particle physicists try to approach the problem from the other end by using precision laboratory experiments situated deep underground. The trouble for these folks lies in the first two letters of the acronym WIMP: these particles are extremely difficult to detect. The putative signal given off when WIMPs interact with normal matter in the laboratory is exceedingly feeble, and it can easily be masked by the background from natural radioactivity and cosmic rays, even in the most exquisitely sensitive of experiments. For this reason, no direct signal of WIMP interactions has yet been confirmed.

Enter Chang and colleagues¹ with the latest report. Using a high-altitude balloon-borne detector called ATIC (for advanced thin ionization calorimeter), they have detected a significant bump in the smooth, background

spectrum of galactic electrons, confirming earlier hints⁸ of its existence. This feature is located at an energy of about 620 GeV (1 GeV is 10⁹ electronvolts and, by mass–energy equivalence, corresponds roughly to the mass of a proton). And it is consistent with the type of signal expected when KK WIMP particles interact and annihilate into electron–positron pairs. (The positron is the electron’s anti-particle: essentially, an electron but with a positive charge.) The process basically amounts to two KK particles disappearing from ‘the dark side’ and appearing in our realm in the form of an electron–positron pair. The signature electrons can then be detected and measured.

What makes the ATIC detection especially intriguing is that the 620-GeV energy of the peak is roughly the mass of KK WIMPs expected from particle physics theories⁷. But in a way, the intensity of the signal is almost too high. To explain its strength requires a large enhancement of local dark matter, such that the Solar System would be whizzing through (or at least near) an especially dense clump of dark matter.

But could there be other explanations for the bump ATIC detected in the electron spectrum? It is certainly possible that known astrophysical objects, such as nearby supernova remnants, spinning pulsars⁹ or, possibly,

microquasars¹⁰ — whether catalogued or not — are responsible for the feature around 620 GeV. And let's not forget that a completely new type of astrophysical object could also produce the detected electron excess; after all, pulsars were discovered only in 1967, and until 1992 we were blissfully unaware of microquasars.

However, all is not lost in our attempts to determine the real source of the ATIC electron feature. Luckily, the expected spectral signature of electrons from nearby KK WIMP annihilation is very distinctive: it shows a gradual rise to a sharp peak and a 'cliff-like' drop-off at the high-energy end (see Fig. 4 on page 364). With a longer observation, or a bigger detector, which would markedly increase the number of electrons detected, one ought to be able to say with some confidence whether the ATIC feature is more like a regular bump (for example, from an astrophysical source) or the sharp, discontinuous feature expected from nearby dark-matter KK WIMP annihilation. In fact, existing γ -ray telescopes, both on the ground and in space, should also be able to search for this possible discontinuity in the spectra of integrated diffuse emission they detect.

In the near future, neutrino telescopes on the scale of cubic kilometres, such as IceCube, buried deep beneath the cold, clear Antarctic ice, could also aid our sleuthing for dark matter¹¹. And we may even have our very own dark-matter factory soon: the Large Hadron Collider is due to start regular operations in the next year or so, and its data will be combed through thoroughly in search of any dark-matter signals.

Squeezing the dark-matter problem from both sides is certainly the way to go: astrophysicists staring up and particle physicists working deep underground may finally unveil the secrets of the dark side. And if it turns out that the ATIC detection really is due to an elusive WIMP about 600 times the mass of a proton, it will go to show just how truly insignificant we are — the wimpy, overweight dark side may have us beaten. ■

Yousaf M. Butt is at the National Academy of Sciences, Washington DC, on leave from the Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA.
e-mail: ybutt@cfa.harvard.edu

- Chang, J. *et al. Nature* **456**, 362–365 (2008).
- Massey, R. *et al. Nature* **445**, 286–290 (2007).
- Spergel, D. N. *et al. Astrophys. J. Suppl. Ser.* **170**, 377–408 (2007).
- Milgrom, M. Preprint at <http://lanl.arxiv.org/abs/0801.3133> (2008).
- McGaugh, S. *Science* **317**, 607–608 (2007).
- Disney, M. J. *et al. Nature* **455**, 1082–1084 (2008).
- Cheng, H.-C. *et al. Phys. Rev. Lett.* **89**, 211301 (2002).
- Torii, S. *et al. Preprint at* <http://lanl.arxiv.org/abs/0809.0760> (2008).
- Hooper, D., Blasi, P. & Serpico, P. D. Preprint at <http://lanl.arxiv.org/abs/0810.1527> (2008).
- Heinz, S. & Sunyaev, R. *Astron. Astrophys.* **390**, 751–766 (2002).
- Hooper, D. & Baltz, E. A. *Annu. Rev. Nucl. Part. Sci.* **58**, 293–314 (2008).

DNA SEQUENCING

Mammoth genomics

Michael Hofreiter

Reconstruction of most of the genome sequence of the woolly mammoth illustrates how such investigations will pave the way for a deeper understanding of the biology and evolution of extinct species.

How times change. Within the space of less than a decade, the development of high-throughput technologies has transformed the task of sequencing a mammalian genome from the years-long, multimillion-dollar endeavour it was originally, to a project that can be performed by an individual laboratory within a few months. Such is the breathtaking extent of progress that it has now been possible to sequence the nuclear genome of the extinct woolly mammoth (Fig. 1) almost to completion.

This achievement is described by Miller *et al.* on page 387 of this issue¹. Stretches of mammoth DNA from a cellular organelle, the mitochondrion, which has its own small genome, have been sequenced previously. But tackling the 4 billion to 5 billion base pairs in the much larger and biologically informative mammoth cell nucleus posed a different order of challenge.

Ancient DNA — DNA obtained from fossils up to about 100,000 years old — is highly fragmented, is present in only trace amounts, and is usually swamped by bacterial and fungal DNA. So the idea of sequencing the complete genome of an extinct species was long unthinkable, given that for 30 years the only large-scale method available, the Sanger sequencing method², was not adequate for such a task. However, in 2005 a new way of sequencing DNA was published³. Known as the 454 method, it initially had a throughput

of 20 million base pairs per run — an increase of two to three orders of magnitude compared with Sanger sequencing. Since then, this method's throughput has increased to about 100 million base pairs and three other types of sequencing approaches have been launched — Solexa, SOLiD and HeliScope. These methods produce as much as 10 billion base pairs of sequence data in a single run. But all of them are 'shotgun' methods, which produce rather short sequencing reads; some also have quite high error rates. So, to obtain a reliable sequence for each nucleotide position, a genome needs to be sequenced to multiple coverage, which Miller and colleagues estimate would be around 10–20-fold for the mammoth. Given the rapid increase in sequencing throughput, it is possible that even this will soon be achieved.

These developments have both changed the way we think about molecular genetics and shown how we might finally get a handle on the genomes of extinct species. Not least, some of the drawbacks of the new technologies for studying modern DNA are advantages when it comes to ancient DNA. Thus, whereas the Sanger method allows individual stretches of up to 800 base pairs to be determined in a single reaction, the new techniques yield much shorter sequences, sometimes as few as 30 base pairs⁴. This is a disadvantage when dealing with modern DNA. But it does not matter with ancient DNA, which is mostly fragmented



Figure 1 | Ancient blueprint. This wall painting of a mammoth and ibex is in Rouffignac Cave, Dordogne, France, and is dated to around 13,000 years ago.

ANCIENT ART & ARCHITECTURE COLLECTION LTD

into pieces shorter than 100 base pairs.

Researchers were quick to exploit the new possibilities. Only months after the 454 technology became available, it was applied to mammoth genomics in a paper⁵ that reported 13 million base pairs of sequence — about 1,000 times more than were covered in the first ancient-genomics study with Sanger sequencing⁶. In that paper⁵, published in January 2006, the authors also announced their plan to sequence the mammoth genome to completion. Miller and colleagues¹ now describe about 70% of the mammoth genome, and so go a long way to achieving that goal.

Miller *et al.* were aided immensely in their task by the fact that, unusually for extinct organisms, some specimens of woolly mammoths have been frozen in permafrost. This is an ideal setting for preserving DNA, and, moreover, for preserving hair, which is an ideal source of DNA for sequencing ancient genomes. If hair still contains DNA, almost all of it will belong to the extinct species, and will not be of bacterial or fungal origin, as is often the case with bones. Thus, the authors needed to sequence a total of 'only' 4.1 billion base pairs to obtain about 3.3 billion base pairs of mammoth DNA. They calculate that the total mammoth genome, estimated at some 4.7 billion base pairs, would have been 1.4 times bigger than the human genome.

Although the mammoth genome is larger than the human genome, the DNA substitution rate seems to be smaller — this is the rate at which one nucleotide replaces another, and so is a measure of evolutionary change. The mammoth genome differs from that of its close relative the African elephant by as little as 0.6%. This is about half the difference between human and chimpanzee, although the two elephant species diverged at about the same time as human and chimpanzee, and probably even slightly earlier (see Fig. 3 of the paper¹ on page 389). For some reason, the substitution rate in the nuclear genome of elephants is much lower than in humans and great apes, a result mirrored in the mitochondrial genomes⁷, where humans and great apes also show a substitution rate more than twice as high as that in the elephant species. As nuclear and mitochondrial genomes are replicated by different enzymes, it remains unclear why both genomes evolve more slowly in elephants than in humans and great apes.

The draft mammoth genome sequence is too fragmented and error-prone to allow standard gene prediction. Nonetheless, Miller and colleagues identified several protein-coding positions that are unique to the mammoth compared with 50 other vertebrate species. The presence of such mammoth-specific differences is not surprising: it is to be expected that each mammalian species contains unique amino-acid substitutions compared with a limited number of other species. For example, the 52-amino-acid fragment of the ATP2C1 protein not only contains an amino-acid

substitution unique to the mammoth, but a further two that are unique to the tenrec and the two-toed sloth, respectively. Similarly, the position in the 30-amino-acid fragment of the protein C1orf190, at which the mammoth differs from other placental mammals, also has amino-acid substitutions in the ground squirrel and kangaroo rat. Although Miller and colleagues argue that the amino-acid differences they identify have a "significantly enhanced likelihood of causing ... phenotypic effects", their analyses by no means prove that an amino-acid substitution has functional consequences or adaptive value. Such questions can be answered only by investigations of the proteins in question.

So what do we learn from the mammoth genome, except that sequencing of complete genomes from extinct species is indeed possible and that there are differences in their DNA sequences compared with those of living animals? As with many draft genome projects, not that much. But a draft genome is only the beginning of the story. The main feature of genome projects is to provide a resource for further research, as vividly shown by the thousands of times the initial human-genome sequencing papers^{8,9} have been cited.

The next draft nuclear genome of an extinct species likely to become available is that of our

closest relative, the Neanderthal, following on from publication of a complete Neanderthal mitochondrial genome sequence¹⁰. For some time yet, much work in genomics will consist of fully annotating and completing genome sequences, as indeed most published sequences of extant vertebrates, let alone that of the extinct mammoth, remain drafts. But when we look further into the future, the task will be to understand which differences at the sequence level underlie the phenotypic differences between a mammoth and an elephant, or a human and a Neanderthal, for which well-annotated genomes provide the essential basis. ■

Michael Hofreiter is at the Max Planck Institute for Evolutionary Anthropology, Deutscher Platz 6, D-04103 Leipzig, Germany.
e-mail: hofreite@eva.mpg.de

1. Miller, W. *et al.* *Nature* **456**, 387–390 (2008).
2. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. Natl Acad. Sci. USA* **74**, 5463–5467 (1977).
3. Margulies, M. *et al.* *Nature* **437**, 376–380 (2005).
4. Smith, D. R. *et al.* *Genome Res.* **18**, 1638–1642 (2008).
5. Poinar, H. N. *et al.* *Science* **311**, 392–394 (2006).
6. Noonan, J. P. *et al.* *Science* **309**, 597–599 (2005).
7. Rohland, N. *et al.* *PLoS Biol.* **5**, e207 (2007).
8. International Human Genome Sequencing Consortium *Nature* **409**, 860–921 (2001).
9. Venter, J. C. *et al.* *Science* **291**, 1304–1351 (2001).
10. Green, R. E. *et al.* *Cell* **134**, 416–426 (2008).

PALAEOCLIMATE

Greenhouse-gas fingerprints

Thomas F. Stocker and Adrian Schilt

Short episodes of warming and cooling occurred throughout the last glaciation. An innovative modelling study indicates that ocean-circulation changes produced much of the causative variation in greenhouse gases.

Much of what we know about abrupt climate change and tipping points in the climate system comes from polar ice cores¹. But these unique data archives provide only a narrow view of the richness of climate dynamics and impacts. Moreover, the origin of the variations in the greenhouse gases associated with the pronounced climate swings during the last ice age, an interval between about 110,000 and 10,000 years ago, remains largely unknown. Hence the significance of computer models in providing a wider perspective.

On page 373 of this issue, Schmittner and Galbraith² present climate-model simulations for an episode of abrupt climate change during the last ice age. Their results show agreement with the palaeoclimatic record³, not only in terms of physical climate variables, but also, remarkably, in changes in the greenhouse gases carbon dioxide (CO₂) and nitrous oxide (N₂O). The researchers conclude that the interaction of physical and biogeochemical processes in the ocean is largely

responsible for the observed variations.

Schmittner and Galbraith² use a coupled climate model of intermediate complexity suitable for palaeoclimatic studies. The physical part of the simulation features a comprehensive ocean-circulation model coupled to an energy-balance model of the atmosphere. A marine-ecosystem module, which includes two classes of phytoplankton, simulates the distribution of nitrate, phosphate, oxygen, inorganic carbon and alkalinity in the ocean. Although the model accounts for carbon cycling in the ocean, atmosphere and terrestrial vegetation, it deals with the global nitrogen cycle only in a simplified fashion. Production of marine N₂O is diagnosed from oxygen concentrations, and the stratospheric sink and the soil source of N₂O are assumed to remain constant.

During the last ice age, the climate system exhibited a series of rapid changes known as Dansgaard–Oeschger events. These involved temperature changes of up to 15 °C in Greenland within a few decades⁴. There is good

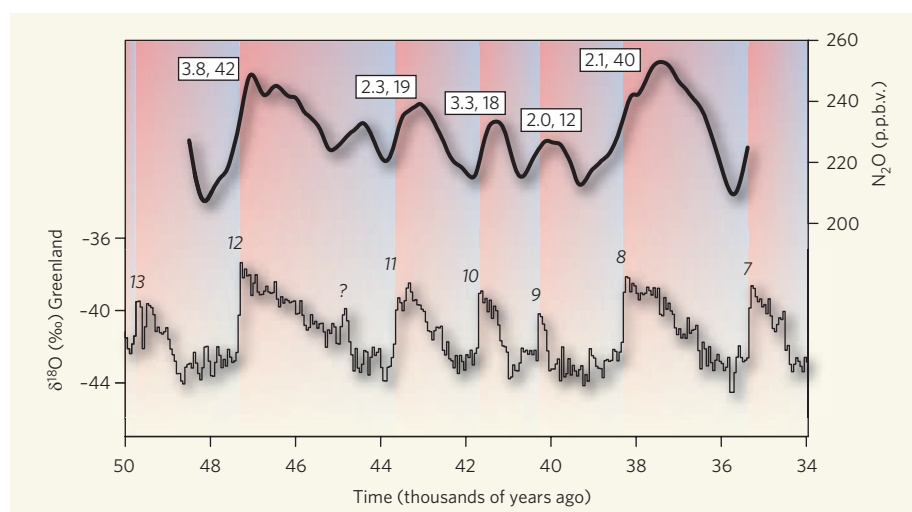


Figure 1 | Records of atmospheric N_2O and of Dansgaard–Oeschger events over a 16,000-year interval during the last ice age. The rich structure of N_2O variations^{3,13} (shown as a spline through the data) correlates well with the incidence of Dansgaard–Oeschger events, the sequence of abrupt warmings recorded in a Greenland ice core plotted here as variation in a proxy, oxygen-isotope measurement, $\delta^{18}O$. Paired numbers in boxes indicate the rate of N_2O increase in p.p.b.v. (parts per billion by volume) per century and the total N_2O increase in p.p.b.v., respectively. Italic numbers denote Dansgaard–Oeschger events in this particular oxygen-isotope record¹⁴. From their modelling work, Schmittner and Galbraith² conclude that the N_2O variations are caused by changes in ocean circulation, which, in turn, influence oxygen concentration and productivity in the ocean.

evidence for the view that a major part in these events was played by the Atlantic meridional overturning circulation (AMOC), the flow of warm surface water to the far north that is balanced by the southward flow of cold water at depth. But as long as the smoking gun is missing, scientists must continue to collect circumstantial evidence from palaeoclimatic data and from model simulations to test this hypothesis. Modellers still do not have the all-encompassing climate model that would simulate a series of Dansgaard–Oeschger events in a self-contained way. They therefore need to resort to provoking such abrupt change in their models by adding and extracting large amounts of fresh water to and from the North Atlantic. The manipulation has the effect of switching the AMOC off and on, and this is what Schmittner and Galbraith do with their climate model.

The physical changes simulated by the model are in reasonable agreement with the palaeoclimatic record, although the amplitude of the temperature change in Greenland is significantly underestimated. This is probably due to insufficient sea-ice response, which is known to amplify temperature changes in the northern North Atlantic⁵. At the time when a reduced AMOC is responsible for cold temperatures in Greenland, temperatures in Antarctica start rising slowly, a phenomenon referred to as the thermal bipolar seesaw⁶. The Antarctic warming correlates strongly with the increase in the concentration of atmospheric CO_2 . In this model, the increase is primarily caused by a reduction in the efficiency of nutrient use by phytoplankton in surface waters of the Southern Ocean — and hence, when they die, of the transport of carbon from the surface to depth. The response is physically driven by

more deep-water formation in the Southern Ocean when deep water in the North Atlantic recedes during a shutdown of the AMOC.

A notable difference between the model results and the palaeoclimatic data appears during the end of the Antarctic warming. In the temperature records, this change extends over about 2,000 years. The model's north–south coupling seems too strong, in that the abrupt resumption of the AMOC causes a similarly abrupt cooling in Antarctica. This may be due to the absence of a dynamical response in the atmosphere, and Schmittner and Galbraith's decision to keep wind patterns constant⁷.

The model does not capture the full complexity of the marine nitrogen cycle⁸; rather, a simple empirical relationship between oxygen content and changes in marine N_2O production is used. So a combination of local mixing, water-mass distribution and variability in the carbon cycle determines the concentrations of atmospheric N_2O . Also, the N_2O contribution from soils, which is known to be important, is held constant. In spite of these pragmatic simplifications, the simulations exhibit good agreement with the palaeoclimatic record, not only in amplitude but also in temporal behaviour — at least in the first phase of the abrupt climate change.

Shutdown of the model AMOC causes a cooling in the Northern Hemisphere, which results in better ventilation of the surface waters in the eastern parts of the equatorial Pacific and Indian oceans. Better ventilation increases oxygen content, thereby reducing N_2O emissions from the ocean and hence N_2O concentrations in the atmosphere. The model also seems to capture a characteristic fingerprint of N_2O changes during Dansgaard–Oeschger events

first found in ice-core data³: longer coolings produce larger reductions in N_2O . This good agreement leaves little room for an effect from N_2O emissions on land, surprisingly so given that about two-thirds of the N_2O emissions today come from the terrestrial biosphere.

Overall then, Schmittner and Galbraith's model² performs remarkably well. But there is obvious room for improvement. For example, the simulation of atmospheric N_2O is poor after the abrupt cooling event has ended. The ice-core record clearly shows a peak of N_2O followed by a slow decrease that evolves over several centuries in the sequence of Dansgaard–Oeschger events (Fig. 1). Neither feature is captured by the model. As the major areas of N_2O production are located in the eastern ocean basins, where upwelling of nutrient-rich water directly responds to the wind, one wonders whether changes in wind patterns and strength, which must have occurred during Dansgaard–Oeschger events⁹, might have been responsible for the rapid N_2O increase at the time of abrupt warming. Also, it is inferred from the ice-core methane (CH_4) record that the water cycle changed significantly during Dansgaard–Oeschger events¹⁰ because the main source of CH_4 is wetlands. Therefore, a better account of soils and their changes will not only permit the simulation of CH_4 , but may also contribute to an improved understanding of the centennial variations in N_2O that have not yet been captured in simulations.

Further progress will come with the generation of more data from isotopic studies of palaeoclimatic archives. Changes in the stable-isotope concentrations of greenhouse gases such as ^{13}C (CH_4), 2H (CH_4), ^{15}N (N_2O) and ^{18}O (N_2O) are powerful fingerprints of different factors affecting the climate system and show their response to climate change. Because such measurements are so challenging, only a fraction of the information that is still locked up in polar ice cores has yet been revealed^{11,12}. There is thus a unique opportunity for those working with physical–biogeochemical models such as that of Schmittner and Galbraith. Before more — and more detailed — isotopic data from polar ice cores become available, modellers must venture predictions of the climate changes that these data will reveal. This will be the least prejudiced approach to the problem — and so the best test-bed for our understanding of how oceanic, atmospheric and biogeochemical processes operate and interact in the climate system. ■

Thomas F. Stocker and Adrian Schilt are in the Division of Climate and Environmental Physics, and the Oeschger Centre for Climate Change Research, University of Bern, Sidlerstrasse 5, 3012 Bern, Switzerland.

e-mail: stocker@climate.unibe.ch

1. EPICA Community Members. *Nature* **444**, 195–198 (2006).
2. Schmittner, A. & Galbraith, E. D. *Nature* **456**, 373–376 (2008).
3. Flückiger, J. et al. *Glob. Biogeochem. Cycles* doi:10.1029/2003GB002122 (2004).

4. Huber, C. *et al.* *Earth Planet. Sci. Lett.* **243**, 504–519 (2006).
5. Flückiger, J. *et al.* *Clim. Dyn.* **31**, 633–645 (2008).
6. Stocker, T. F. & Johnsen, S. J. *Paleoceanography* doi:10.1029/2003PA000920 (2003).
7. Levermann, A., Schewe, J. & Montoya, M. *Geophys. Res. Lett.* doi:10.1029/2007GL030255 (2007).
8. Gruber, N. & Galloway, J. N. *Nature* **451**, 293–296 (2008).
9. Timmermann, A. *et al.* *J. Clim.* **20**, 4899–4919 (2007).
10. Dällenbach, A. *et al.* *Geophys. Res. Lett.* **27**, 1005–1008 (2000).
11. Sowers, T., Alley, R. B. & Jubenville, J. *Science* **301**, 945–948 (2003).
12. Fischer, H. *et al.* *Nature* **452**, 864–867 (2008).
13. Flückiger, J. *et al.* *Science* **285**, 227–230 (1999).
14. North GRIP Members. *Nature* **431**, 147–151 (2004).

CELL BIOLOGY

Nuclear order out of chaos

Tom Misteli

How cells build their internal structures remains one of the central mysteries in cell biology. If the cell nucleus is anything to go by, stochastic assembly and self-organization seem to be key.

As anyone knows who has ever put together one of those home-assembly bookshelves, most man-made structures can only be built by following a defined sequence of steps. It is difficult to set aside this preconceived notion of linear assembly when thinking about how cellular structures emerge. Yet a landmark paper by Kaiser, Intine and Dundr¹, published in *Science*, makes a powerful case for the formation of biological structures by apparently random pathways and self-organization.

The mammalian cell nucleus is a prototypical, highly organized biological structure². Not only does it hold most of an individual's genetic information, but it is also the site of many essential processes, such as reading, copying and repairing the genome. To coordinate these events — and presumably to make them more efficient — the nucleus is highly compartmentalized and contains numerous nuclear bodies that have distinct functions².

Kaiser and colleagues¹ set out to investigate the formation of a prominent nuclear compartment called the Cajal body, a roughly spherical structure of 0.5–1 micrometres diameter. There are typically several Cajal bodies in every nucleus, often near clusters of genes encoding histone proteins or small nuclear RNAs (snRNAs)³. Their precise function is unclear, but the fact that they contain factors involved in small-RNA processing suggests they might be involved in modifying nuclear RNAs and in recycling RNA-processing factors³. Like all other nuclear compartments, Cajal bodies lack a membrane, and are highly dynamic yet stable steady-state structures, exchanging their proteins rapidly and continuously with the surrounding nuclear space².

How Cajal bodies, and indeed all other nuclear bodies, are formed and maintained without membrane boundaries has been a puzzle. Two fundamentally different models have been considered⁴. One idea — akin to the bookshelf-assembly strategy — is that Cajal-body proteins bind to scaffold proteins in an orderly manner to gradually build up the structure. Indeed, scaffold

function has been ascribed to two prominent Cajal-body marker proteins, coilin and SMN (Fig. 1a). Alternatively, it has been suggested that nuclear bodies form through self-organization, whereby their components simply aggregate in a stochastic, largely random way (Fig. 1b).

To distinguish between these two models, Kaiser *et al.*¹ first asked a fundamental question: can Cajal bodies form *de novo*? The authors used a bacterial tethering system to immobilize individual Cajal-body proteins at an engineered random site in the genome of mammalian cells. Not entirely unexpectedly, coilin and SMN were each sufficient to form an apparently normal, functional Cajal body. Surprisingly, however, when the authors tethered minor Cajal-body components to the same genomic site, just about any Cajal-body component could initiate the formation of this structure, including all RNA-processing factors. In fact, minor components were more efficient at forming Cajal bodies than coilin and SMN.

These findings lead to two main conclusions. First, the formation of Cajal bodies does not require a specific gene locus, and could potentially occur anywhere in the genome. Second, this process does not require a strict hierarchical

pathway, as any Cajal-body protein can initiate the formation of the entire structure.

These observations do not agree well with a linear-assembly model and point to self-organization as a driving force in the assembly. Several other features of Cajal bodies also favour a self-organization model. First, Kaiser and colleagues¹ show that the recruitment of any native protein to a newly formed Cajal body occurs with similar kinetics, suggesting largely random assembly. Second, previous work has shown that coilin and SMN can self-assemble⁵. This tendency of the two proteins might be crucial for the assembly process, as Kaiser *et al.* show that Cajal bodies form less efficiently in the absence of either coilin or SMN, hinting that these marker proteins serve to stabilize transient interactions among other Cajal-body components. Third, Cajal-body proteins are highly dynamic, and move rapidly between this structure and its surroundings — a requirement for a self-organization process. But perhaps most telling is the authors' finding¹ that, despite the ready availability of building blocks, the size of the structure formed *de novo* is limited, a classic hallmark of self-organization.

Are these results relevant to other nuclear structures? Probably. For one thing, Kaiser and colleagues report that tethering of PML — a marker protein of another nuclear body — leads to the formation of a PML body. Furthermore, much of what we know about nuclear bodies suggests shared principles of organization. Like the Cajal body, all known nuclear bodies represent dynamic steady-state systems that undergo rapid exchange of components². Moreover, other nuclear bodies also seem to form at specific nuclear sites. Whereas Cajal bodies preferentially associate with clustered histone and snRNA genes⁶, the nucleolus — the most prominent nuclear substructure — forms around sites of ribosomal-RNA transcription, and compartments enriched in factors mediating pre-messenger-RNA splicing form near the sites where RNA is transcribed. The likeliest possibility, therefore, is that nuclear bodies assemble at sites of high gene activity

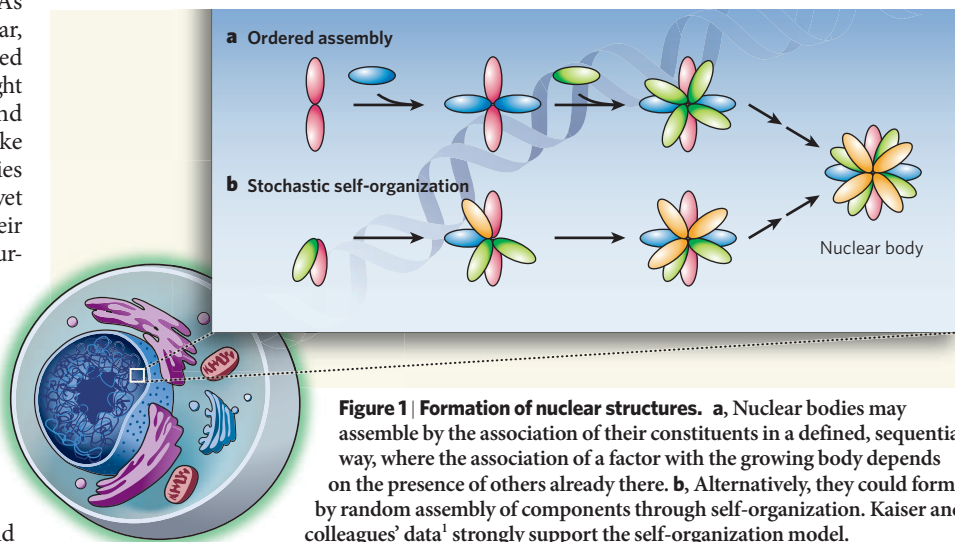


Figure 1 | Formation of nuclear structures. **a**, Nuclear bodies may assemble by the association of their constituents in a defined, sequential way, where the association of a factor with the growing body depends on the presence of others already there. **b**, Alternatively, they could form by random assembly of components through self-organization. Kaiser and colleagues' data¹ strongly support the self-organization model.

through the accumulation of freely diffusing proteins, which leads to the nucleation of a distinct body. In that sense, Kaiser and colleagues' tethering experiments may simply recapitulate the immobilization these factors normally experience at their sites of action.

Having said that, the main caveat of this study¹ is uncertainty about its physiological relevance. It is not known whether *de novo* assembly of Cajal bodies occurs naturally, or, if it does, whether it is a frequent occurrence. In fact, Dundr and colleagues previously observed that, when activated, snRNA gene loci can be recruited to pre-existing Cajal bodies⁷. So it is essential now to establish whether *de novo* assembly or recruitment of pre-assembled Cajal bodies is the preferred means of pairing up these structures with their target genes.

Regardless of this, the present work¹ represents a major step forward. Not only does it provide a conceptual framework for probing nuclear-body formation further, but it also gives us a potential means of succeeding in one of the remaining quests of modern cell biology — finding out how cellular structures form *in vivo*. ■
Tom Misteli is at the National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA.
e-mail: mistelit@mail.nih.gov

NANOTECHNOLOGY

Squaring up with polymers

Anthony J. Ryan

Squares may be unfashionable, but for electronic circuitry no other shape will do. A method for making square arrays of polymeric nanoparticles could herald the next generation of miniature silicon chips.

Just ten years ago, who would have thought that you would be able to carry around your entire music collection in your shirt pocket, or your complete genome on a key ring? These amazing feats depend on miniature information-storage devices, which in turn are based on tiny integrated circuits. Although current technology is incapable of scaling down circuits much further, self-assembling nanoparticles could be pressed into service to make much smaller circuitry elements. But there's a problem — nanoparticles self-assemble into hexagonal arrays that are incompatible with the square arrangements used in industry-standard circuits. Reporting in *Science*, Tang *et al.*¹ describe a method in which appropriately designed polymers form nanoparticles that assemble into square arrays.

The tiny structures found in integrated circuits are currently made by photolithography — a technique that uses ultraviolet light to create patterned masks from light-sensitive polymers, which are then used to control the etching of structures into the surfaces of silicon chips. The resolution of this

technique is limited to about 100 nanometres (the wavelength of the light used to make the masks). But round nanoparticles can self-assemble into smaller patterns, packing together like oranges in a box. If this process could be used to make smaller features on circuits, it would make a huge difference to the amount of data that can be stored on a chip: a tenfold decrease in feature size could increase information density 1,000-fold.

Self-assembled structures are well known in nature — molecules such as lipids form both nanostructures (vesicles and micelles) and thin films (membranes). The synthetic analogues of lipids are block copolymers. Eleven years ago, it was shown² that cheap processes can be devised in which vast arrays of these synthetic polymeric nanoparticles form spontaneously, creating structures ten times smaller than can be made by expensive photolithography. The electronics industry immediately saw the potential benefits of this, but was unable to find a cost-effective way of reworking circuit-design software and fabrication protocols to cope with the hexagonal arrays formed by self-assembly. The Semiconductor Industry

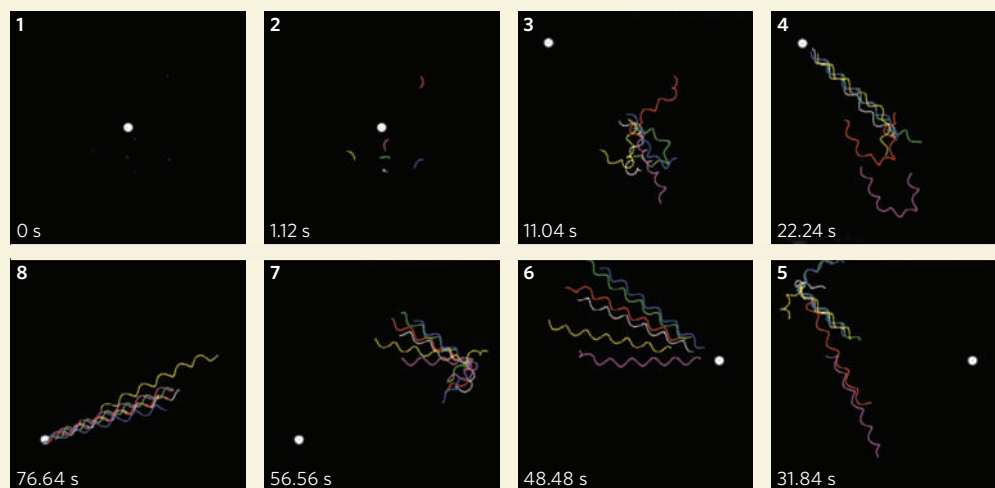
1. Kaiser, T. E., Intine, R. V. & Dundr, M. *Science* doi:10.1126/science.1165216 (2008).
2. Misteli, T. *Cell* **128**, 787–800 (2007).
3. Handwerker, K. E. & Gall, J. *Trends Cell Biol.* **16**, 19–26 (2006).
4. Misteli, T. *J. Cell Biol.* **155**, 181–185 (2001).
5. Hebert, M. D. & Matera, A. G. *Mol. Biol. Cell* **11**, 4159–4171 (2000).
6. Shopland, L. S. *et al. Mol. Biol. Cell* **12**, 565–576 (2001).
7. Dundr, M. *et al. J. Cell Biol.* **179**, 1095–1103 (2007).

CELL BIOLOGY

Why little swimmers take turns

Gáspár Jékely and colleagues' paper on page 395 literally takes the eye. It provides a neat dissection of how the two eyespots in the larva of a marine worm (*Platynereis dumerilii*) perceive light, and then how the larva responds with differential beating of cilia and a directed, helical swimming action (G. Jékely *et al. Nature* **456**, 395–399; 2008).

These stills come from a dynamic computer simulation, created from first principles, in the authors' Supplementary Information. They show the movement of six virtual larvae that are attracted to a shifting light source; the simulation is initiated in frames 1 and 2, and shows the larval reaction when the light source moves to three different positions (frames 3, 4; then 5, 6; and finally 7, 8). The spiral trajectories traced by each of the organisms closely mirror the behaviour



of real *Platynereis* larvae.

This apparently inefficient form of motion, conclude Jékely *et al.*, is optimal for combining eyespot light detection with precise navigation. The authors also point out that at

least some of the same principles probably apply to the many members of the zooplankton that swim in spiral fashion.

More about the paper can be found in 'Making the paper'

(page xiii), and 'An eye for the eye' (page 304) provides a pictorial gallery of ocular evolution. The computer simulation can be viewed at <http://tinyurl.com/59ah73>.

Tim Lincoln

F. NEDELEC

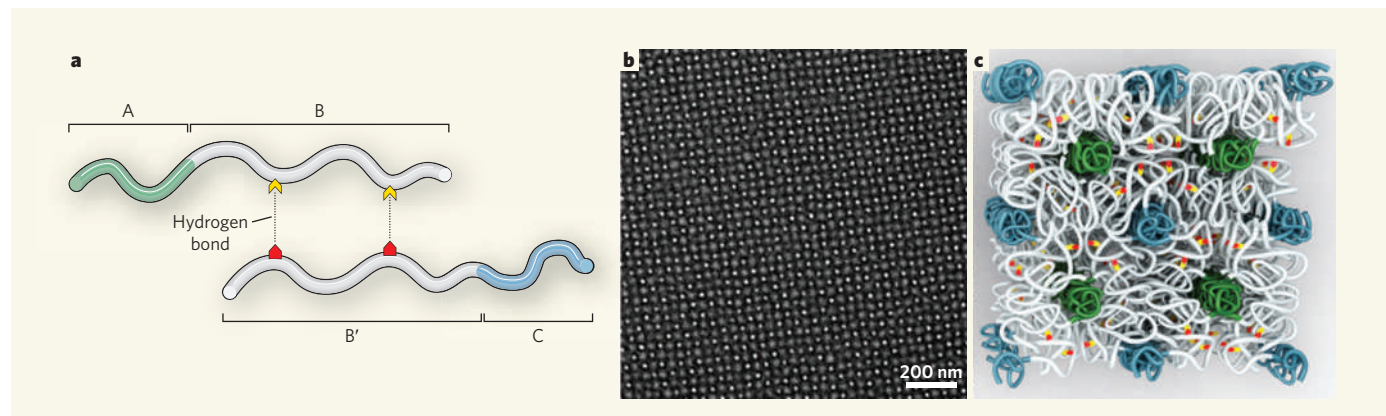


Figure 1 | Square arrays of polymeric nanoparticles. Tang *et al.*¹ have used a blend of diblock copolymers to make a thin film that contains square arrays of cylindrical nanoparticles. **a**, The blend consists of two kinds of diblock copolymer, designated here as A-B and B'-C. The A and C blocks are mutually repulsive and are incompatible with blocks B and B'. Block B contains small numbers of groups (yellow) that form hydrogen bonds to complementary groups (red) in block B'. The two kinds of diblock therefore associate, effectively forming triblock polymers. **b**, This transmission electron microscope image shows that, in a thin film of the blend, blocks A and C form a square lattice of cylindrical nanoparticles (grey and white dots respectively) in a matrix of mixed B and B' blocks (dark regions). **c**, This cartoon reveals how hydrogen bonding between complementary groups (red and yellow) in the B and B' blocks (white) hold the diblock molecules together. Cylindrical domains of A blocks (green) and C blocks (blue) form square arrays. (**b,c**, Courtesy AAAS.)

Association therefore challenged researchers to develop a method for making square arrays of self-assembled block-copolymer particles. Tang *et al.*¹ have risen to this challenge by using a combination of theoretical predictions and methods for controlling non-covalent interactions between molecules.

Like oil and water, most blends of two different polymers split into separate phases because of nonspecific, dispersive interactions between the two types of molecule. But by linking pairs of different molecules (A and B) with a covalent bond (to form an A-B diblock copolymer), one can prevent phase separation into two layers and gain some control over the nanostructures formed as the molecules aggregate. The nanostructures formed by such copolymers are the result of competition between molecular chain-stretching and interfacial effects around the covalent bond connecting the two polymer domains.

Much is known³ about the phase behaviour of diblock copolymers. In bulk material, the conditions needed to form three-dimensional cubic structures with long-range order are well established. In thin films deposited on substrates, where the thickness of the film is similar to the spacing between nanoparticles in the arrays, the preferred patterns are: hexagonal arrays of spheres; stripes of cylinders of block A embedded in a matrix of block B (with the stripes laid parallel to the film's surface); or single bilayers (or trilayers) of nanoparticles. The patterns observed depend on the composition of the copolymers and the interactions of the polymer molecules with the substrate and with the air above the film³.

But more complex molecules, such as A-B-C triblock copolymers — which contain three types of polymer-chain linked in series — can adopt a host of geometries inaccessible to simple A-B diblocks, including chiral and tetragonal arrays formed from cylindrical nanoparticles. Most importantly for forming patterned

masks on circuits, A-B-C copolymers form stable, tetragonal arrays of cylinders in thin films. Moreover, the cylinders align perpendicular to the film's surface, unlike A-B diblocks, in which cylinders normally lie parallel to the surface.

There are, however, several problems associated with A-B-C triblocks. Not only are they more difficult to synthesize, but they also suffer from greater packing 'frustration' than do simple A-B diblocks, because the different blocks in each molecule prefer to pack together with blocks of the same type from other molecules. Domains of each block-type thus form, which means that each molecule must pass through two interfaces between different domains.

So although certain triblock compositions can make square arrays in thin films, long-range order is suppressed. Many of these issues can be resolved by making polymer blends of different diblock copolymers (A-B/B'-C or A-B/C-D), which provide the advantages of having more than two distinct polymer blocks while avoiding the problems associated with connecting all the blocks in one molecule. But it has proved impossible to achieve long-range order in such blends because of their overwhelming propensity to generate separate phases, just as homopolymers do.

Tang *et al.*¹ overcome this problem by using hydrogen bonding between a pair of A-B/B'-C diblocks. In their blend, blocks A and C are chemically different, and both are mutually incompatible with the B and B' blocks. The authors' neat trick is to introduce small amounts of complementary hydrogen-bonding groups into the B and B' blocks (Fig. 1a), so that the blocks form hydrogen bonds to each other. The resulting complex of diblocks behaves like a triblock, and phase separation of the diblocks is suppressed because the hydrogen bonding overcomes the nonspecific dispersive interactions.

The authors made their diblock copolymers using state-of-the-art methods to accurately control both the molecular weight of each block and the number of incorporated hydrogen-bonding groups. They then prepared the A-B/B'-C blend simply by mixing solutions of the two diblocks, and created thin films of the blend (about 50 nanometres thick) on silicon wafers. Using atomic force microscopy (AFM) to observe the nanoscale structure of the films, the authors showed that cylindrical nanoparticles in the film formed perpendicular, square arrays with long-range order (over an area of about 5 × 5 micrometres; Fig. 1b).

They obtained the best results when nearly equal numbers of complementary hydrogen-bonding units were incorporated into the B and B' blocks. Blends that lacked complementary hydrogen-bonding groups formed only small nanostructured regions, which had hexagonal local order. The lack of nanostructure formation in such blends is probably caused by phase separation of the diblocks.

Tang and colleagues' numerical simulations¹ of the packing in their blend reveal that square-packed cylinders have a lower free energy than hexagonally packed cylinders, unlike pure A-B diblocks. This is because a square lattice allows a more uniform distribution of C blocks around each A block (and vice versa), which minimizes unfavourable stretching of the B blocks even though the cylinders are not packed as closely as they could be. Although the A and C units are not connected on the same chain, dynamic hydrogen bonding ensures that, on average, an A-B diblock is always connected to at least one B'-C diblock. This favours B-B' mixing at uniform chain-stretching, so that the blend mimics the behaviour of A-B-C triblocks.

The authors used their block-copolymer films as lithographic masks to transfer a template image into a silicon substrate. Using

scanning electron microscopy, they observed that cylindrical pores (with diameters of about 22 nanometres) were formed in the substrate, spaced 50 nanometres apart. This was consistent with the patterns seen by AFM in the freshly formed films, and demonstrated a high degree of fidelity in the pattern-transfer process.

Could Tang and colleagues' approach be used in the mass production of integrated circuits? IBM already use block copolymers as masks for photolithography to make hexagonal arrays of cylindrical pores in insulators, so it seems that industry could easily adopt the authors' method. But one inherent problem is

that of residual structural defects, especially when using large films — the larger the area covered by a mask, the more likely it is that a defect will occur on the underlying wafer. One solution could be to make mosaics of films from submillimetre-sized patches of polymers, so that the discovery of a defect doesn't require the whole wafer to be rejected; instead, any device using the wafer can be programmed to avoid using the defective 'tile'.

A combination of Tang and colleagues' technique with the latest technologies for making patterned diblock copolymers on patches and in channels could be used to make arrays of

magnets that store trillions of bits of information per square centimetre. But this would be only the start: further developments will undoubtedly lead to greater miniaturization of silicon chips, and to the creation of electronic devices powerful beyond our imagination. ■

Anthony J. Ryan is at the Faculty of Science, University of Sheffield, Sheffield S3 7HF, UK. e-mail: tony.ryan@sheffield.ac.uk

1. Tang, C., Lennon, E. M., Fredrickson, G. H., Kramer, E. J. & Hawker, C. J. *Science* **322**, 429–432 (2008).
2. Park, M., Harrison, C., Chaikin, P. M., Register, R. A. & Adamson, D. H. *Science* **276**, 1401–1404 (1997).
3. Russell, T. *Curr. Opin. Colloid Interface Sci.* **1**, 107–115 (1996).

PHYSIOLOGY

Courier service for ammonia

Mark A. Knepper

Physiological studies in mice demonstrate a surprising role for a kidney protein related to the rhesus factor of red blood cells. Similar research would aid further annotation of mammalian genomes.

The completion of the Human Genome Project spawned a new era of biological research, with many declaring that we have now entered the 'post-genomic' era. This declaration is premature, not least because comprehensive functional annotations are lacking for most of the 20,000-odd protein-coding genes in the human genome¹. Without this knowledge, we cannot gain a thorough understanding of human disease. So a crucial task ahead is to discover the *in vivo* function of each gene product. By investigating the function of the protein encoded by a gene called *Rhcg*, using classical physiological techniques in mice, Biver and colleagues² (page 339 of this issue) provide a model for how this goal can be effectively pursued.

We are constantly facing an acid threat. When oxidized in the body, excess dietary proteins produce sulphuric and phosphoric acids, posing the danger of a lethal drop in blood pH. But kidneys prevent our demise by removing these acids through the excretion of hydrogen ions (H^+). Most of this H^+ is excreted in the urine as ammonium (NH_4^+), which is made by cells in the proximal tubule of the kidney from the amino acid, glutamine.

From the proximal tubule, a series of NH_4^+ -transport processes, ending with a transport step across the cells of the collecting duct, ensure the transfer of NH_4^+ to the urine³ (Fig. 1). In the final step, NH_4^+ is believed to enter the collecting-duct cells by direct transport⁴, but leaves these cells by the

parallel transport of its constituents — H^+ and ammonia (NH_3) — into the urine⁵. An ion pump called V-ATPase mediates H^+ transport. As for the ammonia, because of its small size (17 daltons), it was believed to diffuse freely across the membranes of collecting-duct epithelial cells into the urine, independently of a transporter protein⁵. But is this really the

case? Biver *et al.*² suspected otherwise, and so tested the hypothesis that the protein product of *Rhcg* is responsible for ammonia transport out of collecting-duct cells.

The protein in question, Rhcg, is related to the rhesus (Rh) antigen protein of red blood cells that is used to determine immunological compatibility before blood transfusions. It had previously been shown⁶ that the amino-acid sequences of mammalian Rh-family proteins are similar to those of known ammonia-transport proteins of bacteria, fungi, plants and invertebrates. Indeed, when investigators expressed mammalian Rhcg in cells that do not normally express it, they detected transport of ammonia⁷. But this observation is not a definitive proof for the physiological role of Rhcg at its natural levels in mammalian tissues. What's more, when *Rhbg* — the gene for another Rh-related protein of the collecting duct — was deleted in mice, no measurable defects were detected in ammonia transport in the kidneys⁸.

Undaunted by the negative result with *Rhbg*, Biver *et al.*² deleted the *Rhcg* gene in mice. The authors demonstrate that, contrary to the free-diffusion hypothesis⁵, Rhcg absence results in a two-thirds reduction in the rate of ammonia movement across the collecting-duct epithelial cells. This indicates that, at most, only one-third of ammonia crosses the lipid membrane by diffusion. So Rhcg seems to function as a membrane channel, allowing the passage of ammonia across the lipid membrane between collecting-duct cells and the urine (Fig. 1). The absence of this channel in the *Rhcg*-knockout mice results in decreased excretion of ammonia in the urine and a marked drop in blood pH, clearly establishing a role for Rhcg in the regulation of pH in body fluids.

Biver and colleagues also report a reduction in reproductive capability of male *Rhcg*-knockout mice, associated with altered epididymal function. Like the kidney's collecting duct, the

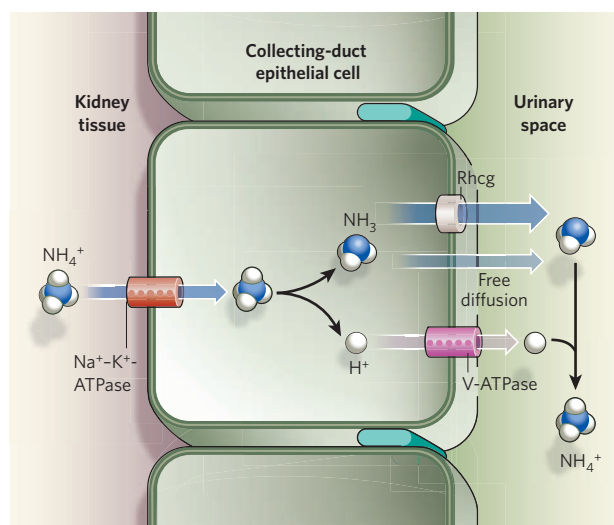


Figure 1 | Ammonium excretion. Acids generated from the metabolism of excess dietary protein are excreted largely as urinary ammonium (NH_4^+) produced by cells in the proximal tubule of the kidney. NH_4^+ is ultimately transported into the urine from collecting-duct cells through parallel movement of hydrogen ions (H^+) and ammonia (NH_3). Biver *et al.*² show that not all ammonia moves by free diffusion — as was thought previously — and that most of it crosses through the Rhcg protein, which functions as an ammonia channel. The V-ATPase pump mediates H^+ transport into the urine, where it recombines with NH_3 to form NH_4^+ . As for the initial NH_4^+ entry into these cells, an ion pump called $Na^+-K^+-ATPase$, which can carry NH_4^+ in lieu of potassium, has been proposed to be involved⁴.

epididymis — a convoluted tube within the testes where sperm maturation occurs — acidifies its lumen using the V-ATPase H^+ pump⁹. Here, however, acidification has a different role: the acidic environment is necessary for sperm maturation and for making sperm cells dormant until they are 're-awoken' by the alkaline environment of the seminal vesicle. In this context, excessive ammonia secretion into the epididymal lumen would be detrimental, prematurely awakening sperm cells. The authors propose that Rhcg in the epididymis could be responsible for scavenging the potentially toxic ammonia from the lumen and transporting it into the cells lining the epididymal lumen where it can be converted to glutamine.

Trying to identify functional differences between knockout and normal mice is technically challenging, in part because of the small size of mice. What helped Biver *et al.* in their quest was their ability to adapt some classical physiological techniques to the mouse, including measurements of the composition of body fluids, of intracellular pH using fluorescent dyes, and of solute transport in isolated perfused renal tubules. These approaches have been part of the physiological armamentarium for years, but have gradually been displaced by modern molecular and cell-biological techniques. For example, the isolated perfused tubule technique developed¹⁰ in the 1960s — like many other intricate physiological methods — has not been used extensively since the 1980s. With expertise in many physiological techniques in decline, Biver and colleagues' work is a testament that such methods are still of enormous value. It will be interesting to see

whether physiologists' new role as functional annotators of the genome will spur a resurgence of training in traditional physiological techniques.

Biver and colleagues' data² resolve a long-standing question in kidney physiology: what is the molecular basis of ammonia transport into the urine? With the identification of Rhcg as the ammonia channel in the collecting duct, researchers can now address questions such as whether and how this channel may be regulated, and what its involvement is in disease processes associated with impaired regulation of blood pH. Reproductive biologists will undoubtedly soon be investigating the specific role of Rhcg in epididymal function and sperm maturation. ■

Mark A. Knepper is in the Division of Intramural Research, National Heart, Lung and Blood Institute, National Institutes of Health, Bethesda, Maryland 20892-1603, USA.
e-mail: knep@helix.nih.gov

1. Clamp, M. *et al.* *Proc. Natl Acad. Sci. USA* **104**, 19428–19433 (2007).
2. Biver, S. *et al.* *Nature* **456**, 339–343 (2008).
3. Knepper, M. A., Packer, R. & Good, D. W. *Physiol. Rev.* **69**, 179–249 (1989).
4. Wall, S. M., Davis, B. S., Hassell, K. A., Mehta, P. & Park, S. J. *Am. J. Physiol. Renal Physiol.* **277**, F866–F874 (1999).
5. Pitts, R. F. *Fed. Proc.* **7**, 418–426 (1948).
6. Marini, A.-M., Urrestarazu, A., Beauwens, R. & André, B. *Trends Biochem. Sci.* **22**, 460–461 (1997).
7. Westhoff, C. M. & Wylie, D. E. *Transfus. Clin. Biol.* **13**, 132–138 (2006).
8. Chambrey, R. *et al.* *Am. J. Physiol. Renal Physiol.* **289**, F1281–F1290 (2005).
9. Pastor-Soler, N., Piétrement, C. & Breton, S. *Physiology* **20**, 417–428 (2005).
10. Burg, M., Grantham, J., Abramow, M. & Orloff, J. *Am. J. Physiol.* **210**, 1293–1298 (1966).

STRUCTURAL BIOLOGY

Enzyme knocked for a loop

Ronald L. Mellgren

Protein-digesting enzymes are kept on a tight leash to stop them from wantonly attacking targets. Two crystal structures show how an inhibitory protein domain gags one such enzyme without being chewed up itself.

All organisms have enzymes, known as proteinases, that break down other proteins, and have vital roles in such disparate processes as food digestion and blood clotting. These enzymes can't be allowed to run amok — if a protein has been mistakenly broken down, then a replacement has to be painstakingly re-synthesized from scratch. Several strategies have therefore evolved to ensure that proteinases are allowed to act only in appropriate circumstances. For example, inhibitor proteins block proteinase activity using various mechanisms^{1,2}. In this issue, two crystallographic studies^{3,4} show that the protein inhibitor calpastatin uses a previously undiscovered mechanism to block the activity of the calpain

proteinase. Two helical regions of calpastatin bind to widely separated sites on calpain; this reinforces the wrapping of an intervening, unstructured region of the inhibitor around calpain's active site.

Calpains are calcium-dependent proteinases found inside cells⁵. In most mammalian cells, the two main forms are μ -calpain (calpain 1) and m-calpain (calpain 2), each of which consists of two protein chains: a large catalytic subunit and a smaller regulatory subunit. It has been shown that m-calpain is required for embryonic development in mice⁶. Calpains also cut several structural proteins that regulate cell shape, disrupting their association with other proteins and so affecting the organization of



50 YEARS AGO

Suspect Documents: Their Scientific Examination. By Dr. Wilson R. Harrison — The book under review is a very important contribution to the published knowledge on the examination of suspect documents ... Fraudulent documents have existed as long as genuine documents, and the problem of their detection has been of considerable importance to States and individuals. This problem existed before science was sufficiently advanced to assist in its solution and, therefore, there grew up a tradition of examination of documents based on superstitious credulity and with no scientific background whatsoever ... [T]he volume under review fills the gap from the British point of view ... and the book must therefore be read by all who are interested in these matters.

ALSO:

The control of diseases of livestock is an important feature in the general economy of most countries in improving animal production and in fostering international trade in livestock and animal products ... It is gratifying to learn in this connexion that one of the recent steps taken ... has been the establishment of a World Reference Laboratory for Foot-and-Mouth Disease at the Research Institute (Animal Virus Diseases), Pirbright, Surrey.
From *Nature* 22 November 1958.

100 YEARS AGO

A common criticism of the methods of teaching science adopted in schools for girls is that they are too academic and have little or no bearing upon the duties the girls will be called upon to perform in after life. This weakness is, we are glad to know, becoming less common, and earnest efforts are being made in several centres to arrange courses of work in which elementary science and the home arts are taught together, the latter being treated largely as applications of the former.
From *Nature* 19 November 1908.

50 & 100 YEARS AGO

the cell's protein 'skeleton' (the cytoskeleton). Studies on calpain-deficient cells show that calpains are important for remodelling the cytoskeletal structures that abut the cell membrane and regulate cell movement⁷. Furthermore, calpains contribute to several calcium-initiated cell-death programs.

Calpastatin inhibits both μ - and m-calpains, and is approximately the same size as the large calpain subunits. But unlike the proteinases, it is an intrinsically unstructured protein, adopting a defined structure only on binding to active calpain. Each calpastatin molecule contains four regions — the calpastatin inhibitory domains, CIDs I–IV — that bind calpains and block their activity (Fig. 1a). Each CID is, in turn, subdivided into three regions (A to C) that are predicted to interact with calpain. Because calpastatin molecules are largely unravelled in solution, they can potentially interact with four calpain molecules simultaneously, with one calpain at each CID.

The current papers^{3,4} reveal for the first time a complete picture of how calpastatin shuts down calpain activity. Hanna *et al.*³ (page 409) present the structure of calcium-bound m-calpain in complex with the CID-IV fragment of calpastatin, whereas Moldoveanu *et al.*⁴ (page 404) report the structure of calcium-bound m-calpain in complex with the CID-I fragment. The two studies show that the binding of calcium ions to m-calpain causes changes in the positioning of the four catalytic subunit domains DI–DIV, and of a regulatory subunit domain, DVI (Fig. 1b). Calcium binding thus generates interaction sites in the proteinase to which CIDs bind; specifically, a site created in DIV is bound by CID region A, and a site in DVI is bound by CID region C. In this way, CIDs bind simultaneously to two widely separated domains of calpain. This strong binding allows CID region B — a largely unstructured area between the helices of regions A and C — to make several contacts with the DI–DIII regions of calpain, thus blocking the proteinase's active site. These observations^{3,4} confirm and refine the findings of recent nuclear magnetic resonance studies⁸, which also indicated that calpastatin wraps around calpain as described above.

This inhibitory mode of action by calpastatin seems like a dangerous trick — deliberately inserting an unfolded protein into the active site of a proteinase is a bit like putting one's head in a lion's mouth and hoping that it won't be bitten off. But the crystal structures^{3,4} show that calpastatin has evolved a neat trick to avoid being cleaved by calpain: the short stretch of protein that would be expected to enter the active site actually loops away from the proteinase (Fig. 1b). Thus, calpastatin, although largely disordered on its own, is fine-tuned to generate a local structural motif that protects it from attack when attached to its calcium-bound substrate.

The new studies^{3,4} tell us as much about calpains as they do about calpastatin. Although

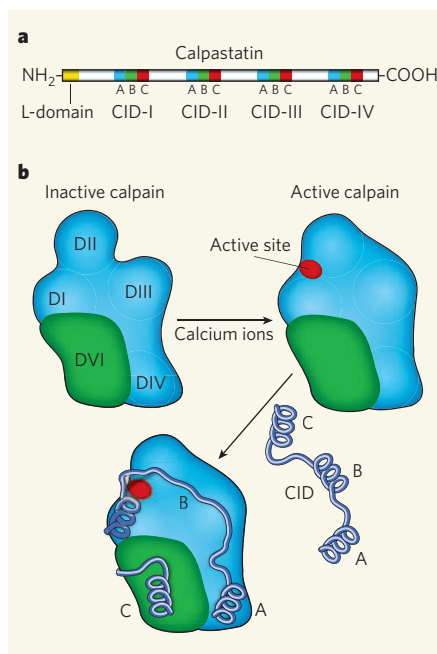


Figure 1 | It's a wrap for calpain. Calpain enzymes break down proteins, but are inhibited by calpastatin. **a**, Calpastatin is an elongated protein that contains three or four independent inhibitory domains (CIDs). The larger form of calpastatin (which contains four CIDs) also has a terminal L-domain that anchors the protein to cell membranes. Each CID is composed of three regions, A–C, which bind to different domains of calpain. **b**, Hanna *et al.*³ and Moldoveanu *et al.*⁴ report crystal structures of a calpain (m-calpain) in complex with calcium ions and with CID fragments of calpastatin. These structures show calpain's four catalytic subunit domains (DI to DIV, blue) and its regulatory subunit domain DVI (green). When m-calpain binds to calcium, it adopts a compact, active conformation in which DIII moves toward DII, and the active site (red) is formed. The A region of the CID associates with DIV, whereas the C region binds to DVI. The B region wraps around the remaining subunits, blocking the active site. The B region loops away from the catalytic site to avoid being cleaved by the enzyme.

the crystal structure of calcium-free (and thus inactive) m-calpain was solved several years ago⁹, crystallization of calcium-bound calpain has been problematic. Structural studies have thus relied on the extrapolation of results from engineered calpain fragments that contain only DI and DII (ref. 10). Calpastatin does not bind calcium, and has little affinity for calpains in the absence of calcium. It therefore seems highly likely that, as the authors of the studies suggest^{3,4}, the structure of calcium-bound calpain in their calpastatin co-crystals is at least a close approximation of the active conformation of m-calpain.

The structures^{3,4} show that, on calcium binding, four specific arginine amino-acid side chains in the DIII domain of m-calpain contact the DII domain, stabilizing the catalytically active conformation. Moldoveanu *et al.*⁴ observe that mutations in m-calpain, in

which the arginines are replaced with other amino acids, lower m-calpain's activity. These arginines are also found in a muscle-specific isoform of calpain known as calpain-3; mutation of the arginines in calpain-3 leads to loss of the proteinase's activity, an effect that has been associated with some cases of limb-girdle muscular dystrophy¹¹.

More broadly, the studies^{3,4} provide further evidence for the involvement of m- and μ -calpains in continuous cellular processes. Calpastatin CIDs apparently evolved to remain intact as they inhibit calpain activity, even though resistance to cleavage is not necessary for this purpose³. But if cleaved and reused, CIDs would probably be less effective at blocking calpain activity than the intact inhibitors. It therefore seems that there is an evolutionary advantage in salvaging CIDs for future inhibition cycles, to avoid organisms using vital resources to make new molecules whenever the need arises. It is thus reasonable to conclude that calpains are activated within cells often enough to warrant this energy-saving measure.

Previous studies have shown that calpains can cleave calpastatin in the disordered regions between CIDs, even at relatively high concentrations of calpastatin¹². The resulting calpastatin fragments are themselves calpain inhibitors. Some calpastatin is anchored to membranes through a terminal region (the L-domain, Fig. 1a), but the fragmentation mechanism could provide water-soluble calpain inhibitors that access different subcellular regions. Moldoveanu *et al.*⁴ show that region B of CID-I does not seem to inhibit μ -calpain well — could this explain how apparently inhibited calpain can still find a way to cleave the calpastatin to which it is bound? By damaging calpastatin, calpain could paradoxically sow the seeds of its own defeat, because, once formed, the calpastatin fragments have the potential to effectively shut down the proteinase's activity throughout the cell. With such fascinating nuances still to be explained, the calpain–calpastatin system certainly deserves continued exploration. ■

Ronald L. Mellgren is in the Department of Physiology and Pharmacology, University of Toledo College of Medicine, Toledo, Ohio 43614-2598, USA.
e-mail: ronald.mellgren@utoledo.edu

- Borth, W. *FASEB J.* **6**, 3345–3353 (1992).
- Estrada, S., Olson, S. T., Raub-Segall, E. & Bjork, I. *Protein Sci.* **9**, 2218–2224 (2000).
- Hanna, R. A., Campbell, R. L. & Davies, P. L. *Nature* **456**, 409–412 (2008).
- Moldoveanu, T., Gehring, K. & Green, D. R. *Nature* **456**, 404–408 (2008).
- Goll, D. E., Thompson, V. F., Li, H., Wei, W. & Cong, J. *Physiol. Rev.* **83**, 731–801 (2003).
- Dutt, P. *et al. BMC Dev. Biol.* **6**, doi:10.1186/1471-213x/6/3 (2006).
- Franco, S. J. & Huttenlocher, A. *J. Cell Sci.* **118**, 3829–3838 (2005).
- Kiss, R. *et al. FEBS Lett.* **582**, 2149–2154 (2008).
- Hosfield, C. M., Elce, J. S., Davies, P. L. & Jia, Z. *EMBO J.* **18**, 6880–6889 (1999).
- Moldoveanu, T. *et al. Cell* **108**, 649–660 (2002).
- Jia, Z. *et al. Biophys. J.* **80**, 2590–2596 (2001).
- Mellgren, R. L., Mericle, M. T. & Lane, R. D. *Arch. Biochem. Biophys.* **246**, 233–239 (1986).

A role for Rhesus factor Rhcg in renal ammonium excretion and male fertility

Sophie Biver^{1*}, Hendrica Belge^{2*}, Soline Bourgeois^{3,4*}, Pascale Van Vooren¹, Marta Nowik³, Sophie Scohy^{1†}, Pascal Houillier⁴, Josiane Szpirer¹, Claude Szpirer¹, Carsten A. Wagner³, Olivier Devuyst² & Anna Maria Marini¹

The kidney has an important role in the regulation of acid–base homeostasis. Renal ammonium production and excretion are essential for net acid excretion under basal conditions and during metabolic acidosis. Ammonium is secreted into the urine by the collecting duct, a distal nephron segment where ammonium transport is believed to occur by non-ionic NH₃ diffusion coupled to H⁺ secretion. Here we show that this process is largely dependent on the Rhesus factor Rhcg. Mice lacking Rhcg have abnormal urinary acidification due to impaired ammonium excretion on acid loading—a feature of distal renal tubular acidosis. *In vitro* microperfused collecting ducts of *Rhcg*^{−/−} acid-loaded mice show reduced apical permeability to NH₃ and impaired transepithelial NH₃ transport. Furthermore, Rhcg is localized in epididymal epithelial cells and is required for normal fertility and epididymal fluid pH. We anticipate a critical role for Rhcg in ammonium handling and pH homeostasis both in the kidney and the male reproductive tract.

Ammonium is a principal nitrogen source for microorganisms and plants, whereas in animals it is best known for its cytotoxic effects that may lead to hepatic encephalopathy for instance¹. Because more than 98% of ammonium is in the NH₄⁺ form at physiological pH, throughout we use ‘ammonium’ to refer to the sum of ammonium (NH₄⁺) and ammonia (NH₃) if not otherwise specified. In mammals, the liver and kidney are involved in ammonium detoxification and excretion, respectively. Ammonium metabolism and transport in the kidney are particularly critical for systemic pH regulation^{2–5}. Along the nephron, ammonium, first produced in the proximal tubule, is secreted in the lumen of this segment and then reabsorbed at the level of the thick ascending limb of the loop of Henle. These processes involve nonspecific ammonium transport across several systems in which ammonium competes with natural substrates such as potassium. About 80% of the ammonium produced in the proximal tubule is finally secreted into the urine by the more distal nephron segment, the collecting duct. Since the first description of the concept in the late 1940s, the process of rapid transepithelial transport of ammonium in the collecting duct is thought to occur through non-ionic NH₃ diffusion across the lipid bilayer^{6,7}. NH₃ is then trapped as NH₄⁺ in the lumen of the collecting duct by buffering protons secreted throughout ATPases present at the apical membrane of acid-secreting (type A) intercalated cells. The inability of the distal nephron to appropriately acidify urine in humans results in complete or incomplete forms of distal renal tubular acidosis (dRTA), with potential associated metabolic acidosis reflecting the altered pH homeostasis^{8,9}.

In the meantime, genes for specific ammonium transport proteins were identified by functional complementation in yeast, thereby defining the Mep-Amt family of proteins largely represented in microorganisms, plants and invertebrates but absent from vertebrate animals^{10,11}. Rhesus (Rh) factors have previously been shown to be related to proteins from the Mep-Amt family¹² and could therefore correspond to their, yet to be described, vertebrate counterpart. Although the main Rhesus antigen (RHD) was identified decades

ago¹³, the physiological role of Rhesus-type proteins remains largely unknown. Human Rhesus factors comprise the blood-group antigens (RHCE and RHD)^{14,15}, their associated glycoprotein (RHAG)¹⁶, and two non-erythroid members (RHBG and RHCG)^{17–19}. The Rhb and Rhcg proteins are expressed in distal segments of the mammalian nephron, the connecting tubule and the collecting duct^{20–22}. In the acid-secreting intercalated cells of the mouse kidney, basolateral Rhb coexists with apical Rhcg^{20,21}. Human RHAG and RHCG were shown to mediate bidirectional ammonium transport when expressed in yeast¹⁷. This function of Rhesus proteins in ammonium transport is sustained by several functional expression studies^{23–29}. However, genetic ablation of mouse Rhb yielded no impairment of renal ammonium excretion and no phenotype³⁰. Another potential role for Rhesus proteins, including human RHAG, has been suggested in CO₂ transport^{31–33} but this role is still questioned³⁴, and the biological function of Rhesus factors thus remains debated.

Here we targeted the *Rhcg* gene in mouse and show that the Rhcg protein is required for urinary ammonium excretion by being critical for NH₃ transport across the collecting duct epithelium. Hence, the longstanding assumption that this process occurs solely by lipid phase diffusion of NH₃ should be revised.

Decreased ammoniuria in *Rhcg*^{−/−} mice

To address the physiological role of Rhcg in mouse, we disrupted the *Rhcg* gene by homologous recombination deleting exons 3 to 9 coding for transmembrane segments 4 to 12 (Fig. 1a, b). *Rhcg*^{−/−} mice were born with the expected Mendelian ratio. Polymerase chain reaction with reverse transcription (RT–PCR) analysis demonstrated the complete loss of the 2-kb *Rhcg* transcript in *Rhcg*^{−/−} mice (Fig. 1c) and the lack of Rhcg was confirmed by immunoblotting (Fig. 1d) and immunostaining (Fig. 1e). On a standard diet, *Rhcg*^{−/−} mice showed no obvious abnormalities. Broad-spectrum histological analyses revealed no defect, including in kidney structure and the distribution of cell types in the collecting duct (Fig. 1e, Supplementary Figs 1, 2 and

¹Université Libre de Bruxelles (U.L.B.), Institut de Biologie et de Médecine Moléculaires, Laboratoires de Biologie du Développement et de Physiologie Moléculaire de la Cellule, Rue Professeurs Jeener et Brachet, 12, B-6041 Gosselies, Belgium. ²Université catholique de Louvain (U.C.L.), Division of Nephrology, Avenue Hippocrate, 54, B-1200 Brussels, Belgium. ³Institute of Physiology and Zurich Center for Integrative Human Physiology, University of Zurich, Winterthurerstrasse 190, Zurich CH-8057, Switzerland. ⁴INSERM UMR5 872; Université Pierre et Marie Curie; Université Paris-Descartes; F-75006 Paris, France. †Present address: Delphi Genetics, B-6041 Gosselies, Belgium.

*These authors contributed equally to this work.

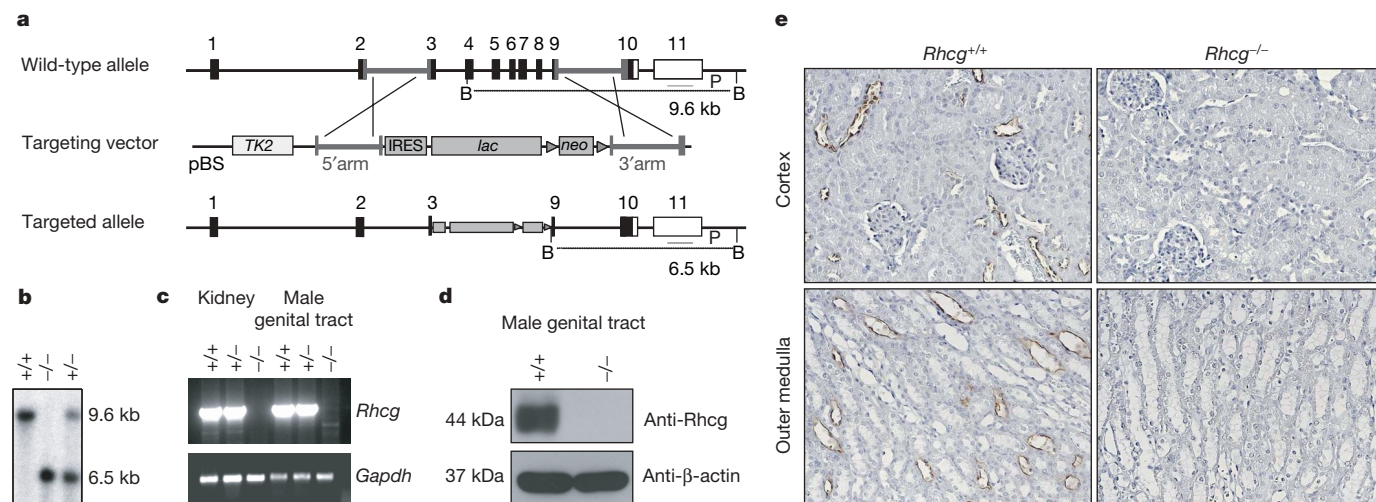


Figure 1 | Gene targeting of *Rhcg*. **a**, The targeting strategy. The IRES-LacZ cassette is the reporter, the pGK-Neo locus is the positive selection marker and thymidine kinase (*TK2*) represents the negative selection marker. The *loxP* sites are shown as triangles, *Rhcg* exons as black boxes. B, BamHI sites used in DNA blot analysis with the exon 11 probe (P). pBS, pBluescript. **b–e**, Demonstration of correct *Rhcg* targeting. **b**, Blot analysis of DNA from littermate progeny of *Rhcg*^{+/-} crosses. BamHI-digested genomic DNA hybridized with the exon 11 probe produces a shorter fragment in *Rhcg*^{-/-}

Supplementary Table 1). However, *Rhcg*^{-/-} mice had a significant decrease in urinary ammonium excretion and a more alkaline urinary pH (Table 1, Supplementary Table 2 and Supplementary Table 3).

Metabolic acidosis in *Rhcg*^{-/-} mice

In humans, defects in distal urinary acidification result in dRTA⁸. Notably, in the incomplete forms of dRTA, the associated metabolic acidosis is not overt but may be revealed by an acid-load challenge. To unmask an incomplete form of dRTA, we investigated the response of

Table 1 | Baseline parameters in *Rhcg*^{-/-} and wild-type mice

	<i>Rhcg</i> ^{+/+}	<i>Rhcg</i> ^{-/-}
Body weight (g)	43 ± 1 (35)	38 ± 1 (31)*
Arterial blood		
Hematocrit (%)	42 ± 1 (12)	41 ± 2 (11)
pH	7.28 ± 0.01 (12)	7.31 ± 0.01 (11)
pCO ₂ (mm Hg)	39 ± 2 (12)	37 ± 2 (11)
Plasma		
HCO ₃ ⁻ (mM)	18 ± 1 (12)	18 ± 1 (11)
Na ⁺ (mM)	147 ± 1 (12)	146 ± 1 (11)
K ⁺ (mM)	4.3 ± 0.1 (12)	3.8 ± 0.1 (11)*
Cl ⁻ (mM)	120 ± 1 (12)	117 ± 1 (11)
Urea (mg dl ⁻¹)	75 ± 6 (9)	69 ± 5 (8)
Creatinine (mg dl ⁻¹)	0.15 ± 0.01 (9)	0.14 ± 0.01 (8)
Ammonium (μM)	49 ± 5 (10)	47 ± 4 (8)
Urine		
Urinary flow rate (μl min ⁻¹ g ⁻¹ BW)	0.035 ± 0.003 (35)	0.034 ± 0.003 (31)
pH	7.42 ± 0.09 (35)	7.68 ± 0.08 (31)*
Na ⁺ (mM mM ⁻¹ creatinine)	32 ± 2 (35)	33 ± 1 (31)
Cl ⁻ (mM mM ⁻¹ creatinine)	36 ± 2 (34)	38 ± 1 (26)
K ⁺ (mM mM ⁻¹ creatinine)	59 ± 2 (35)	60 ± 2 (31)
Calcium (mg mg ⁻¹ creatinine)	0.064 ± 0.01 (32)	0.056 ± 0.01 (27)
Phosphate (mg mg ⁻¹ creatinine)	1.76 ± 0.13 (31)	1.56 ± 0.17 (21)
Aldosterone (ng mg ⁻¹ creatinine)	0.074 ± 0.005 (7)	0.083 ± 0.006 (7)
Creatinine (mM)	5.80 ± 0.29 (35)	6.01 ± 0.30 (31)
Ammonium (mM mM ⁻¹ creatinine)	2.82 ± 0.22 (35)	2.22 ± 0.17 (30)*

Means ± s.e.m. (n); BW, body weight.
*, *P* < 0.05 versus wild type.

DNA (6.5 kb). **c**, RT-PCR analyses on renal and male genital tract RNA with primers amplifying exons 2–7 of *Rhcg* (control: *Gapdh*). **d**, Immunoblotting analysis. Membrane proteins from male genital tract probed with anti-*Rhcg* antiserum (loading control: β-actin). **e**, *Rhcg* immunodetection in kidney sections (counterstaining: haematoxylin). Apical *Rhcg* (brown) detected in intercalated and principal cells of wild-type collecting ducts. Histological analysis of *Rhcg*^{-/-} kidneys reveals no structural abnormalities. Original magnification, ×320.

Rhcg^{-/-} mice to a short (2 days) and long (6 days) NH₄Cl acid load (Fig. 2a–f, Supplementary Table 2 and Supplementary Table 3). In wild-type mice, the acid load induced a transient decrease of blood pH and bicarbonate levels with partial recovery at day 6 because of adaptation (Fig. 2a, b). In contrast, in *Rhcg*^{-/-} mice, the decrease in blood pH and bicarbonate levels lasted during the entire treatment, indicating an impaired ability to cope with the acid load. *Rhcg*^{-/-} mice also showed lower arterial pCO₂ levels at day 6, suggesting that respiratory compensation for metabolic acidosis occurs (Fig. 2c).

Both wild-type and *Rhcg*^{-/-} animals responded to the acid load by increasing the urinary excretion of ammonium and titratable acidity, resulting in a large increase in net acid excretion (NAE) and a concomitant reduction in urinary pH (Fig. 2d–f). However, NAE was considerably lower in *Rhcg*^{-/-} mice owing to a selective defect in ammoniuria (Fig. 2e, f and Supplementary Table 2). *Rhcg*^{-/-} mice had a ~40% reduction in urinary ammonium excretion, and were unable to maximally acidify urine (Fig. 2d and Supplementary Table 2). Defective ammonium excretion is not related to impaired ammonium synthesis, because blood glutamine and glutamate levels, and messenger RNA abundance of enzymes involved in renal ammonium production were unaffected by *Rhcg* deletion (Supplementary Fig. 3). The impaired handling of an acid load in *Rhcg*^{-/-} mice is thus caused by a decreased ability to excrete acid in the form of NH₄⁺.

Acid load and weight loss in *Rhcg*^{-/-} mice

We also investigated the response of *Rhcg*^{-/-} mice to HCl, a stronger acid challenge which, in contrast to NH₄Cl, does not modify nitrogen intake. The 7-day HCl load led to a large increase in urinary ammonium excretion in wild-type mice, reaching a plateau after 2 days (Fig. 2g and Supplementary Table 3). *Rhcg*^{-/-} mice again showed an impaired ability to excrete ammonium, barely reaching 25% of the wild-type response. Chronic metabolic acidosis has been shown to induce weight loss³⁵. The HCl load induced a slight and transient reduction of body weight in wild-type mice, whereas *Rhcg*^{-/-} mice progressively lost about 25% of their initial body weight (Fig. 2h). A significantly reduced body weight of *Rhcg*^{-/-} mice was already detectable at baseline, and was sustained in *Rhcg*^{-/-} mice drinking NH₄Cl (Supplementary Table 2). Hence, *Rhcg*^{-/-} animals have an impaired ability to maximally acidify urine owing to impaired ammonium excretion and are therefore unable to properly handle

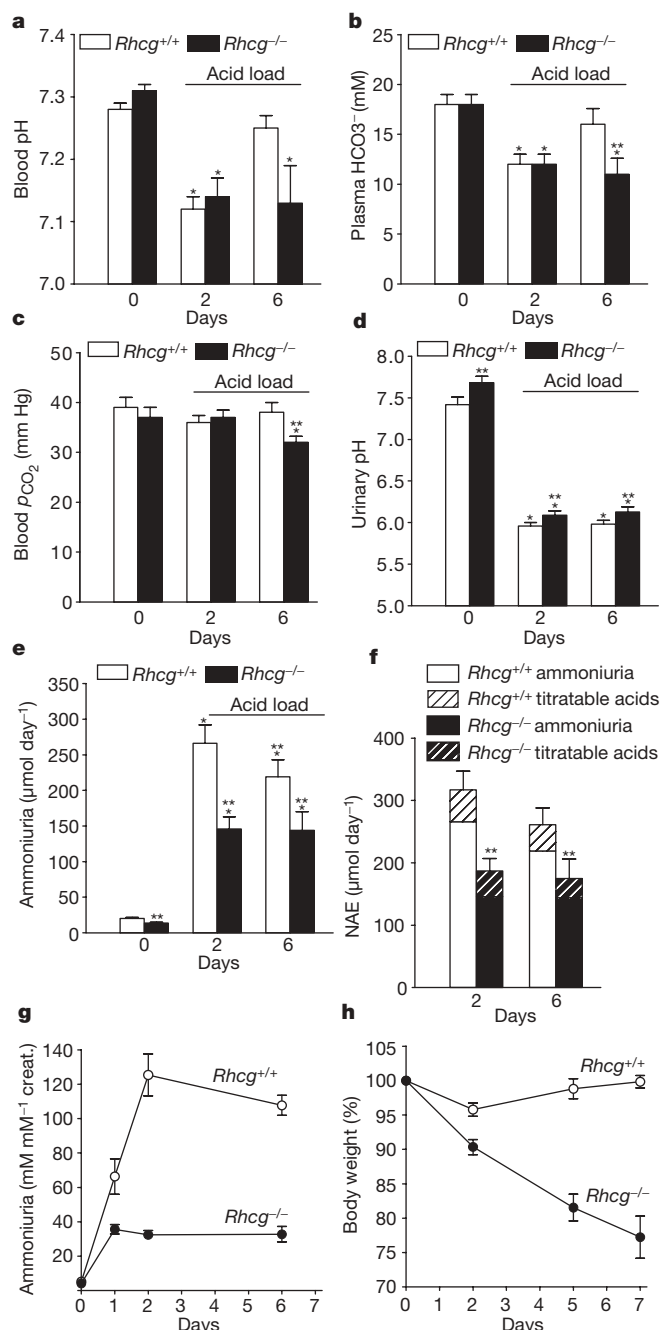


Figure 2 | Impaired acid-stress handling in *Rhcg*^{-/-} mice. **a–f**, Increased metabolic acidosis and impaired acid excretion in *Rhcg*^{-/-} mice (CD1) exposed to an NH_4Cl load for 0, 2 or 6 days. Data expressed as means \pm s.e.m.; $n \geq 11$; asterisk, $P < 0.05$ versus baseline; double asterisk, $P < 0.05$ versus wild type. **a**, Blood pH. **b**, Plasma HCO_3^- . **c**, Blood $p\text{CO}_2$. **d**, Urinary pH. **e**, Ammoniuria. **f**, NAE. The *Rhcg*^{-/-} mice show a strongly reduced NAE owing to a selective defect in ammoniuria, whereas titratable acidity levels are similar to wild type. **g, h**, Impaired ammonium excretion and excessive weight loss in *Rhcg*^{-/-} mice (C57BL/6) exposed to an HCl acid-load. **g**, Time course of urinary ammonium excretion; $n \geq 6$. **h**, Adaptation to the acid-loading challenge indicated by body weight loss, expressed as percentages of initial body weights. HCl-induced weight loss is more pronounced in *Rhcg*^{-/-} mice; $n \geq 4$; creat., creatinine.

acid challenges, demonstrating important characteristics of distal renal tubular acidosis.

Rhcg is required for NH_3 transport

To assess the effect of Rhcg on the apical $\text{NH}_3/\text{NH}_4^+$ permeabilities of collecting duct cells, we first microperfused pH-sensitive dye-loaded

cortical collecting ducts (CCDs) and outer medullary collecting ducts (OMCDs) from *Rhcg*^{-/-} mice after a 2-day acid-load (Fig. 3 and Supplementary Table 4). An inwardly directed ammonium gradient was applied from the lumen (20 mM NH_4Cl ; Fig. 3a). The initial rate of intracellular alkalinization of *Rhcg*^{-/-} CCDs and OMCDs was reduced to 29% and 40%, respectively, compared to wild type (Fig. 3b). As the alkalinization phase is associated with rapid NH_3 entry and subsequent protonation to form intracellular NH_4^+ ³⁶, we conclude that the net permeability to NH_3 is impaired in *Rhcg*^{-/-} mice (Fig. 3b). There was no obvious difference in the subsequent acidification rate, reflecting similar rates of NH_4^+ entry and dissociation of NH_3 and H^+ in wild-type and *Rhcg*^{-/-} OMCD and CCD cells (Fig. 3b). We next directly measured diffusive NH_3 permeability across the CCD epithelium of 2-day acid-challenged mice (Fig. 3c and Supplementary Table 5). Imposing a bath-to-lumen NH_3 gradient, in the nominal absence of a NH_4^+ gradient, generated a measurable NH_3 secretion flux across the CCD epithelium that was reduced to 28% in *Rhcg*^{-/-} compared to wild type. This corresponded to a 67% decrease in NH_3 permeability in *Rhcg*^{-/-} CCDs (Fig. 3c).

Both independent *in vitro* methods thus show Rhcg to be critical for NH_3 transport across the apical membrane of CCDs and allow the identification of the mechanism responsible for reduced urinary ammonium excretion in *Rhcg*^{-/-} mice.

Decreased male fertility in *Rhcg*^{-/-} mice

Because *Rhcg* mRNA has been detected in mouse testis¹⁸, we investigated the fertility of *Rhcg*^{-/-} mice. Although the *Rhcg*^{-/-} males fertilized females as quickly as the *Rhcg*^{+/+} littermates, their litters were ~10% smaller than those of the two other male types (*Rhcg*^{+/-}: 11.7 ± 0.5 versus *Rhcg*^{+/+}: 12.7 ± 0.4 and *Rhcg*^{+/+}: 12.9 ± 0.3 pups per litter, $n \geq 40$, $P < 0.05$). *Rhcg* deletion did not affect the fertility of females. Real-time RT-PCR showed that *Rhcg* is highly expressed in the epididymis and at lower levels in the testis (Fig. 4a). *Rhcg* staining was detected on the luminal membrane of cells lining a subset of epididymal ducts (Fig. 4b). An important characteristic of the epididymal luminal fluid is that it is maintained acidic to keep spermatozoa immotile during their maturation in the epididymis³⁷. Notably, the pH of the epididymal fluid was significantly lower in *Rhcg*^{-/-} mice compared to wild type (Fig. 4c).

Discussion

We have demonstrated that Rhcg is required for maximal urinary acidification and ammonium excretion. This contrasts with the absence of any particular phenotype reported for mice lacking *Rhbh*³⁰. Our data are consistent with a role of Rhcg as an ammonium transport protein mediating the net flux of NH_3 in the collecting ducts. They lead to a new model of renal ammonium excretion occurring at least in part through a protein-mediated pathway, and not only through simple non-ionic diffusion and subsequent luminal trapping as previously suggested^{6,7}. Clearly, a channel/carrier-mediated pathway would provide the kidney with the ability to regulate ammonium excretion at this level. Understanding the physiological function of *Rhbh* and the potential interaction with Rhcg remains a challenge for future investigations.

We have also shown that *Rhcg*^{-/-} mice have features of incomplete human dRTA⁸. Consequently, human *RHCG* may be a candidate gene for inherited forms of dRTA that are not associated with mutations in B1 (*ATP6V1B1*) and a4 (*ATP6V0A4*) subunits of the H^+ V-ATPase or the $\text{HCO}_3^-/\text{Cl}^-$ exchanger AE1 (*SLC4A1*)³⁸. The reduced male fertility and lower pH of epididymal fluid in *Rhcg*^{-/-} mice also indicate that Rhcg has a role in the homeostasis of the male reproductive tract. The impairment of *RHCG* in humans might likewise be accompanied by male fertility problems. Moreover, low ammonium concentrations were reported to inhibit human spermatozoa motility *in vitro*³⁹. Rhcg might thus be involved in ammonium scavenging from the epididymal fluid, consistent with the high

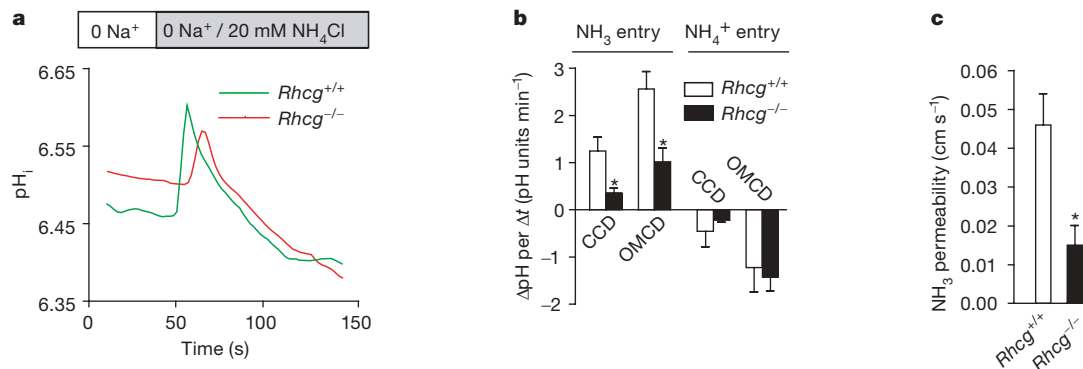


Figure 3 | Reduced NH_3 permeability of microperfused collecting duct segments from $\text{Rhcg}^{-/-}$ acid-challenged mice. NH_4Cl was given for 2 days to CD1 mice. **a**, **b**, CCDs and OMCDs were microperfused *in vitro* in the lumen with 20 mM NH_4Cl . Intracellular pH (pH_i) was monitored with the pH-sensitive dye 2',7'-bis-(2-carboxyethyl)-5-(and -6)-carboxyfluorescein (BCECF) to examine pH_i evolution during luminal NH_4Cl application. **a**, pH_i recording from OMCD cells. **b**, Alkalinization and acidification rates

in CCD and OMCD cells. The reduced rate of alkalinization in $\text{Rhcg}^{-/-}$ cells indicates impaired luminal NH_3 entry in CCD and OMCD cells. Means \pm s.e.m.; $n = 4$ –7 per genotype or nephron segment; asterisk, $P < 0.05$ versus wild type. **c**, Diffusive NH_3 permeability across $\text{Rhcg}^{-/-}$ CCD epithelium, measured by imposing a bath-to-lumen NH_3 gradient in microperfused CCDs, was strongly impaired compared to wild type. Means \pm s.e.m.; $n = 4$; *, $P < 0.05$.

glutamine synthetase activity of the epithelial cells lining the epididymis head⁴⁰ and hence their ability to metabolise ammonium.

By highlighting the role of Rhcg in ammonium disposal and urinary acidification, our study brings new insights in the regulation of acid–base homeostasis, the role of the kidney and the pathophysiology of disease states such as dRTA and male infertility.

METHODS SUMMARY

$\text{Rhcg}^{-/-}$ mice and *in vivo* studies. $\text{Rhcg}^{-/-}$ mice were obtained by homologous recombination, targeting exons 3–9 of Rhcg , resulting in a null allele ($\text{Rhcg}^{\text{tm11bmm}}$). Experiments were performed on adult male $\text{Rhcg}^{+/+}$ and $\text{Rhcg}^{-/-}$ littermates, matched for age. For acid-loading experiments, a solution of 300 mM NH_4Cl (in 2% sucrose) was given to mice in their drinking water for 2 or 6 days. For HCl acid-loading, a 50:50 mixture (w/v) of normal chow and 0.4 M HCl was given for 7 days with water *ad libitum*. All protocols were conducted in accordance with the National Research Council Guide for the Care and Use of Laboratory Animals and were approved by the local Ethics Committee.

Analytic procedures. Urinary parameters were measured on a Synchron CX5 analyser (Beckman). Blood gases and electrolytes were measured using an ABL77 pH/ blood-gas analyser (Radiometer). Ammonium concentrations were measured by glutamate dehydrogenase determination and by using an ammonium combination electrode (DC218- NH_4 , Mettler). Titratable acids were measured using a DL 50 titrator (Mettler).

Immunostaining and immunoblotting. Immunostaining was performed on paraffin-embedded sections, and specificity was demonstrated on similar knockout tissues. Membrane proteins were obtained from the male genital tract homogenized in 250 mM sucrose, 10 mM Tris-HCl, pH 7.5, and protease inhibitors.

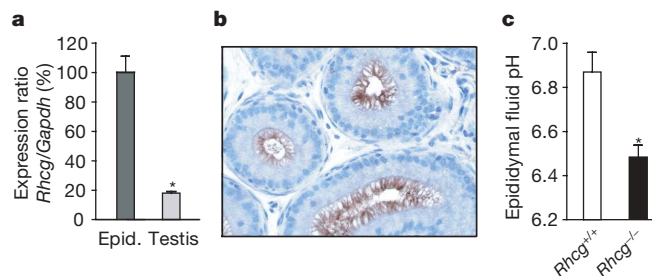


Figure 4 | Rhcg in the male genital tract. **a**, Rhcg expression in wild-type testis and epididymis (epid.) by real-time RT–PCR, epididymal expression taken as 100%. Rhcg is expressed significantly more in epididymis than in testis; $n = 4$. **b**, Rhcg immunodetection in wild-type epididymis (counterstaining: haematoxylin). Original magnification, $\times 200$. Rhcg (brown) is present on the luminal membrane of epididymal cells. **c**, Significant decrease in $\text{Rhcg}^{-/-}$ epididymal fluid pH. Means \pm s.e.m.; $n = 7$; asterisk, $P < 0.05$.

Proteins were separated by SDS–PAGE. The anti- Rhcg antibody was obtained after immunizing rabbits with a carboxy-terminal peptide (EEVNTVYIPEDLAHK) of mouse Rhcg .

***In vitro* microperfusion experiments.** *In vitro* microperfusion of single CCD and OMCD fragments, intracellular pH and transepithelial NH_3 permeability measurements were performed essentially as described previously³⁰.

Statistics. Data are expressed as means \pm s.e.m. Statistical comparisons were tested by Student's *t*-tests using the Graphpad Prism software.

Received 14 June; accepted 29 September 2008.

- Butterworth, R. F. Hepatic encephalopathy. *Alcohol Res. Health* **27**, 240–246 (2003).
- DuBose, T. D. Jr, Good, D. W., Hamm, L. L. & Wall, S. M. Ammonium transport in the kidney: new physiological concepts and their clinical implications. *J. Am. Soc. Nephrol.* **1**, 1193–1203 (1991).
- Hamm, L. L. & Simon, E. E. Roles and mechanisms of urinary buffer excretion. *Am. J. Physiol.* **253**, F595–F605 (1987).
- Knepper, M. A. NH_4^+ transport in the kidney. *Kidney Int., Suppl.* **33**, S95–S102 (1991).
- Weiner, I. D. & Hamm, L. L. Molecular mechanisms of renal ammonia transport. *Annu. Rev. Physiol.* **69**, 317–340 (2007).
- Pitts, R. F. The renal excretion of acid. *Fed. Proc.* **7**, 418–426 (1948).
- Sartorius, O. W., Roemmelt, J. C., Pitts, R. F., Calhoun, D. & Miner, P. The renal regulation of acid–base balance in man. IV. The nature of the renal compensations in ammonium chloride acidosis. *J. Clin. Invest.* **28**, 423–439 (1949).
- Rodríguez Soriano, J. Renal tubular acidosis: the clinical entity. *J. Am. Soc. Nephrol.* **13**, 2160–2170 (2002).
- Laing, C. M., Toye, A. M., Capasso, G. & Unwin, R. J. Renal tubular acidosis: developments in our understanding of the molecular basis. *Int. J. Biochem. Cell Biol.* **37**, 1151–1161 (2005).
- Marini, A. M., Vissers, S., Urrestarazu, A. & Andre, B. Cloning and expression of the *MEP1* gene encoding an ammonium transporter in *Saccharomyces cerevisiae*. *EMBO J.* **13**, 3456–3463 (1994).
- Ninnemann, O., Jauniaux, J. C. & Frommer, W. B. Identification of a high affinity NH_4^+ transporter from plants. *EMBO J.* **13**, 3464–3471 (1994).
- Marini, A. M., Urrestarazu, A., Beauwens, R. & Andre, B. The Rh (rhesus) blood group polypeptides are related to NH_4^+ transporters. *Trends Biochem. Sci.* **22**, 460–461 (1997).
- Levine, P. & Stetson, R. E. An unusual case of intra-group agglutination. *J. Am. Med. Assoc.* **113**, 126–127 (1939).
- Le van Kim, C. *et al.* Molecular cloning and primary structure of the human blood group RhD polypeptide. *Proc. Natl Acad. Sci. USA* **89**, 10925–10929 (1992).
- Cherif-Zahar, B. *et al.* Organization of the gene (*RHCE*) encoding the human blood group RhCcEe antigens and characterization of the promoter region. *Genomics* **19**, 68–74 (1994).
- Ridgwell, K. *et al.* Isolation of cDNA clones for a 50 kDa glycoprotein of the human erythrocyte membrane associated with Rh (rhesus) blood-group antigen expression. *Biochem. J.* **287**, 223–228 (1992).
- Marini, A. M. *et al.* The human Rhesus-associated RhAG protein and a kidney homologue promote ammonium transport in yeast. *Nature Genet.* **26**, 341–344 (2000).
- Liu, Z. *et al.* Characterization of human RhCG and mouse Rhcg as novel nonerythroid Rh glycoprotein homologues predominantly expressed in kidney and testis. *J. Biol. Chem.* **275**, 25641–25651 (2000).

19. Liu, Z., Peng, J., Mo, R., Hui, C. & Huang, C. H. Rh type B glycoprotein is a new member of the Rh superfamily and a putative ammonia transporter in mammals. *J. Biol. Chem.* **276**, 1424–1433 (2001).
20. Quentin, F. *et al.* RhBG and RhCG, the putative ammonia transporters, are expressed in the same cells in the distal nephron. *J. Am. Soc. Nephrol.* **14**, 545–554 (2003).
21. Verlander, J. W. *et al.* Localization of the ammonium transporter proteins RhBG and RhCG in mouse kidney. *Am. J. Physiol. Renal Physiol.* **284**, F323–F337 (2003).
22. Eladari, D. *et al.* Expression of RhCG, a new putative $\text{NH}_3/\text{NH}_4^+$ transporter, along the rat nephron. *J. Am. Soc. Nephrol.* **13**, 1999–2008 (2002).
23. Bakouh, N. *et al.* NH_3 is involved in the NH_4^+ transport induced by the functional expression of the human Rh C glycoprotein. *J. Biol. Chem.* **279**, 15975–15983 (2004).
24. Benjelloun, F. *et al.* Expression of the human erythroid Rh glycoprotein (RhAG) enhances both NH_3 and NH_4^+ transport in HeLa cells. *Pflugers Arch.* **450**, 155–167 (2005).
25. Nakhoul, N. L. *et al.* Characteristics of renal Rhbg as an NH_4^+ transporter. *Am. J. Physiol. Renal Physiol.* **288**, F170–F181 (2005).
26. Zidi-Yahiaoui, N. *et al.* Human Rhesus B and Rhesus C glycoproteins: properties of facilitated ammonium transport in recombinant kidney cells. *Biochem. J.* **391**, 33–40 (2005).
27. Mak, D. O., Dang, B., Weiner, I. D., Foskett, J. K. & Westhoff, C. M. Characterization of ammonia transport by the kidney Rh glycoproteins RhBG and RhCG. *Am. J. Physiol. Renal Physiol.* **290**, F297–F305 (2006).
28. Mayer, M. *et al.* Different transport mechanisms in plant and human AMT/Rh-type ammonium transporters. *J. Gen. Physiol.* **127**, 133–144 (2006).
29. Westhoff, C. M., Ferreri-Jacobia, M., Mak, D. O. & Foskett, J. K. Identification of the erythrocyte Rh blood group glycoprotein as a mammalian ammonium transporter. *J. Biol. Chem.* **277**, 12499–12502 (2002).
30. Chambrey, R. *et al.* Genetic ablation of Rhbg in the mouse does not impair renal ammonium excretion. *Am. J. Physiol. Renal Physiol.* **289**, F1281–F1290 (2005).
31. Endeward, V., Cartron, J. P., Ripoche, P. & Gros, G. RhAG protein of the Rhesus complex is a CO_2 channel in the human red cell membrane. *FASEB J.* **22**, 64–73 (2008).
32. Soupene, E., Inwood, W. & Kustu, S. Lack of the Rhesus protein Rh1 impairs growth of the green alga *Chlamydomonas reinhardtii* at high CO_2 . *Proc. Natl Acad. Sci. USA* **101**, 7787–7792 (2004).
33. Ji, Q. *et al.* CeRh1 (rhr-1) is a dominant Rhesus gene essential for embryonic development and hypodermal function in *Caenorhabditis elegans*. *Proc. Natl Acad. Sci. USA* **103**, 5881–5886 (2006).
34. Missner, A. *et al.* Carbon dioxide transport through membranes. *J. Biol. Chem.* **283**, 25340–25347 (2008).
35. May, R. C., Kelly, R. A. & Mitch, W. E. Metabolic acidosis stimulates protein degradation in rat muscle by a glucocorticoid-dependent mechanism. *J. Clin. Invest.* **77**, 614–621 (1986).
36. Bevensee, M. O. & Boron, W. F. *Control of Intracellular pH* (eds Alpern, R. J. & Hebert, S. C.) 1429–1490 (Elsevier Inc., 2008).
37. Pastor-Soler, N., Pietrement, C. & Breton, S. Role of acid/base transporters in the male reproductive tract and potential consequences of their malfunction. *Physiology (Bethesda)* **20**, 417–428 (2005).
38. Karet, F. E. Inherited distal renal tubular acidosis. *J. Am. Soc. Nephrol.* **13**, 2178–2184 (2002).
39. Kim, S. C. & Kim, H. W. Effects of nitrogenous components of urine on sperm motility: an *in vitro* study. *Int. J. Androl.* **21**, 29–33 (1998).
40. van Straaten, H. W. *et al.* Cellular concentrations of glutamine synthetase in murine organs. *Biochem. Cell Biol.* **84**, 215–231 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We are grateful to P. Drèze, B. André and P. Gabant for their contribution to the initial step of this study. We thank H. Mezdour, L. Pouilly and P. Persoons for microinjection experiments, B. Pajak for help in broad-spectrum histological assessment, V. Beaujean, Y. Cnops, H. Debaix and T. Nishino for excellent technical assistance, and A. Tamah and M.-F. Vincent for amino acid chromatography. These studies were supported by the Belgian agencies FNRS and FRSM (3.4.546.04.F, 3.4.635.05.F, 3.4592.06.F), the ‘Fonds Brachet-recherche’, the ‘Fonds Van Buuren’, the ‘Fondation Alphonse & Jean Forton’, a Concerted Research Action (05/10-328), an Inter-university Attraction Pole (IUAP P6/05), the Swiss National Science Foundation (31-109677), an INSERM grant, and the EuReGene and EUNEFRON (GA#201590) projects of the European Community (FP6 and FP7). S.Bi. was the recipient of a FRIA fellowship; H.B., A.M.M. and C.S. are Research fellow, Research associate and Research director of the FNRS, respectively.

Author Contributions S.Bi., J.S., C.S. and A.M.M. developed the concept of this study; all authors contributed to the intellectual design, execution and interpretation of the experiments; S.Bi. and A.M.M. wrote the paper with specific contribution of all authors and the particular involvement of O.D. and C.A.W. All authors analysed and discussed the data, and commented on the final version of the manuscript. S.Bi., H.B. and S.Bo. share first authorship.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.M.M. (ammardini@dbm.ulb.ac.be).

ARTICLES

Generation of pluripotent stem cells from adult human testis

Sabine Conrad¹, Markus Renninger³, Jörg Hennenlotter³, Tina Wiesner¹, Lothar Just¹, Michael Bonin⁴, Wilhelm Aicher^{5,6}, Hans-Jörg Bühring⁷, Ulrich Mattheus², Andreas Mack², Hans-Joachim Wagner², Stephen Minger⁸, Matthias Matzkies⁹, Michael Reppel⁹, Jürgen Hescheler⁹, Karl-Dietrich Sievert³, Arnulf Stenzl³ & Thomas Skutella^{1,6}

Human primordial germ cells and mouse neonatal and adult germline stem cells are pluripotent and show similar properties to embryonic stem cells. Here we report the successful establishment of human adult germline stem cells derived from spermatogonial cells of adult human testis. Cellular and molecular characterization of these cells revealed many similarities to human embryonic stem cells, and the germline stem cells produced teratomas after transplantation into immunodeficient mice. The human adult germline stem cells differentiated into various types of somatic cells of all three germ layers when grown under conditions used to induce the differentiation of human embryonic stem cells. We conclude that the generation of human adult germline stem cells from testicular biopsies may provide simple and non-controversial access to individual cell-based therapy without the ethical and immunological problems associated with human embryonic stem cells.

The ability to derive pluripotent stem cells from the adult human testis has important implications for biotechnology and regenerative medicine¹. Although these cells are unipotently restricted to the generation of gametes in the course of normal development^{2,3}, several lines of evidence suggest that under certain circumstances, cells of the germ line have the ability to give rise to cells that are pluripotent^{4–6}. The term of pluripotency is differently defined in research with mouse and human stem cells. The NIH and the ISSCR guidelines and criteria for human pluripotency include teratoma formation in addition to microarray assays for transcription factors and other gene activity associated with pluripotency. Teratomas, which are tumours containing different kinds of cells and tissues from all three germ layers at various stages of maturation, occur almost exclusively in the gonads⁷. Furthermore, primordial germ cells (PGCs) give rise to pluripotent cells when cultured under appropriate conditions^{4,8}. PGCs have differentiation properties similar to those of embryonic stem (ES) cells isolated from the inner cell mass⁹. Recently the successful establishment of germline stem cells from neonatal mouse testis was reported⁵. In addition, one study⁶ successfully generated mouse adult germline stem cells (GSCs) with pluripotency from spermatogonial stem cells from adult mouse testis. As in the experiments reported previously⁵, these cells were able to differentiate into derivatives of all germ layers *in vitro*, generated teratomas in immunodeficient mice and, when injected into an early blastocyst, contributed to the development of various organs. Similar results with GPR125⁺ germline progenitor cells have been reported by another study¹⁰.

Generation of pluripotent human adult GSCs

We used in total 22 different human testicular parenchymas to generate human adult GSCs. The obtained tissues were mechanically and enzymatically dissociated and filtered to obtain a single-cell suspension containing cells of varying sizes and shapes (Supplementary Fig. 1a).

In a next step the single cells were cultured for 4 days in uncoated dishes with knockout culture medium with glia-derived neurotrophic factor (GDNF), a growth factor essential for the self-renewing division of spermatogonial stem cells¹¹, or cultured directly in leukaemia inhibitory factor (LIF; ref. 12)-supplemented medium (basic medium), which is sufficient to maintain mouse ES cells or embryonic germ cells^{13,14} in an undifferentiated state (Supplementary Fig. 1a). Under these conditions most of the single testis cells attached to the culture plate. For pre-selection of spermatogonial cells with magnetic-activated cell separation (MACS) we used CD49f (α_6 integrin)^{15,16}, a marker selected by us from the different tested surface antigens. By using other antibodies like CD90 (Thy-1) or GDNFR- α 1, which have been described in the literature for mouse spermatogonial stem cell enrichment (see for example refs 17, 18), or CD133 (a marker for human ES and precursor cells) we achieved comparable but not better selection (data not shown). An important tool to gain a highly pure spermatogonial cell population is the subsequent matrix selection procedure with collagen and laminin¹⁹ to extract spermatogonial cells for further cultivation with basic medium and LIF to generate human adult GSCs. With this procedure we were able to obtain a pure population of spermatogonial cells (VASA⁺ vimentin[−]) and completely deplete somatic cells (VASA[−] vimentin⁺) (Fig. 1). After this selection and purification, colonies of spermatogonial cells appeared (Fig. 2a, panel 1) and increased in size (Fig. 2a, panel 2). After 10–15 days, these colonies changed their morphology (Fig. 2a, panel 3), became multilayered and clearly demarcated colonies with boundaries appeared (Fig. 2a, panel 4). These colonies continued to increase in number and size (Fig. 2a, far right panel). Functional proof is provided by the fact that the negative fraction of somatic cells (VASA[−] vimentin⁺) did not form stem cell colonies under LIF supplementation and were negative for stem cells markers, and even more importantly did not form any teratomas (Supplementary Fig. 1). In contrast, even after long-term

¹Institute of Anatomy, Department of Experimental Embryology, ²Institute of Anatomy, Department of Cellular Neurobiology, Österbergstraße 3, 72074 Tübingen, Germany.

³Department of Urology, University Clinic Tübingen, Hoppe-Seyler-Straße 3, Tübingen 72076, Germany. ⁴Institute of Anthropology and Human Genetics, Microarray Facility, University Clinic, Calwerstraße 7, 72076 Tübingen, Germany. ⁵ZMF Research Laboratories, University Clinic Tübingen, Waldhörnlestraße 22, 72072 Tübingen, Germany. ⁶Center for Regenerative Biology and Medicine (ZRM), Paul-Ehrlich-Straße 15, 72076 Tübingen, Germany. ⁷Department of Internal Medicine II, University Clinic Tübingen, Otfried-Müller-Straße 10, 72076 Tübingen, Germany. ⁸Stem Cell Biology Laboratory, Wolfson Centre for Age-Related Diseases, King's College London, King's College, London SE1 1UL, UK. ⁹Institute of Neurophysiology, University of Cologne, Robert-Koch-Straße 39, 50931 Cologne, Germany.

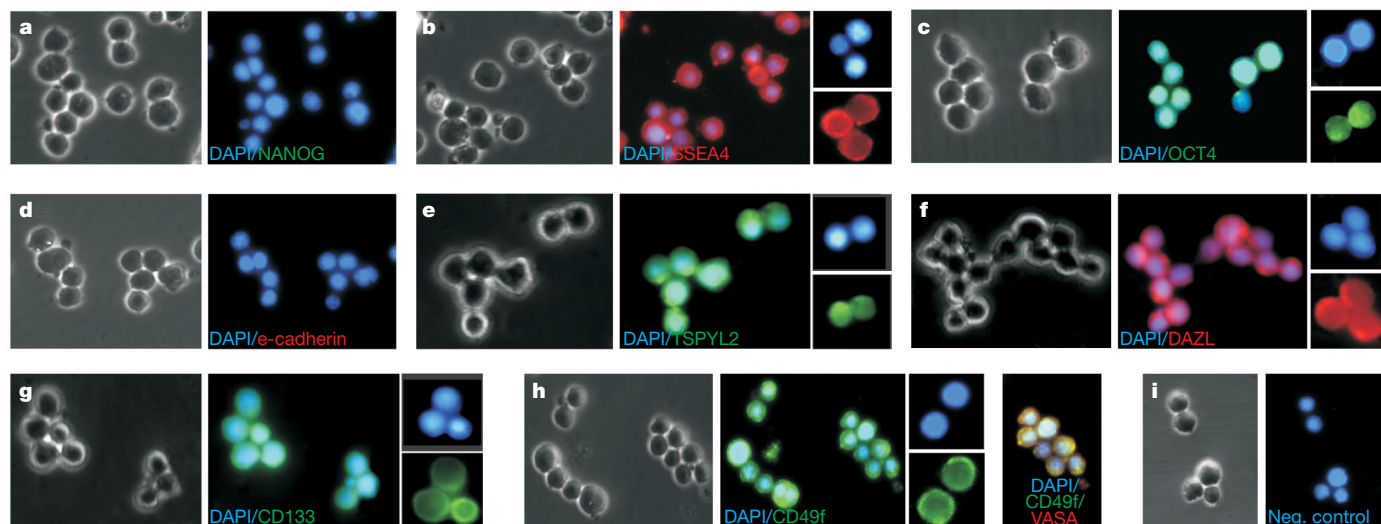


Figure 1 | Selection of spermatogonial cells from adult human testis.

a–i, Immunohistochemistry with the germ and stem cell markers NANOG, SSEA4, OCT4, e-cadherin, TSPYL2, DAZL, CD133 and the selection marker CD49f in spermatogonial cells. Co-expression of CD49f and VASA in

spermatogonial cells is shown. Nuclei were stained with 4,6-diamidino-2-phenylindole (DAPI) and magnifications (original magnification $\times 40$) show nuclear or cytoplasmic staining.

cultivation the VASA⁺ vimentin[−] human adult GSC colonies behaved more like human ES cells in their molecular profile and differentiation capacity, and formation of teratomas.

Electron microscopy of purified spermatogonial cells showed typical morphological characteristics of spermatogonia (Supplementary Fig. 1b). A more detailed immunohistochemical characterization revealed that these cells were positive for VASA, SSEA4, OCT4, TSPYL2, DAZL, CD133 and CD49f, but negative for NANOG and e-cadherin and the somatic marker vimentin (Figs 1, 3a and Supplementary Fig. 1a).

Generally the cells were passed 3 to 4 weeks after initiation of the culture. After passing the multilayered colonies reappeared with constant doubling times up to high passages (over 40 and higher).

We also tested different medium conditions for their ability to induce the generation of human adult GSCs (Fig. 2d). To do this, we plated spermatogonial cells for 14 days under different conditions and observed the formed colonies over another 42-day period (Fig. 2d). Without supplementation with GDNF and LIF no clusters were formed at all. In contrast, both LIF alone and GDNF followed by LIF resulted in a constant rate of cluster formation over 42 days; with

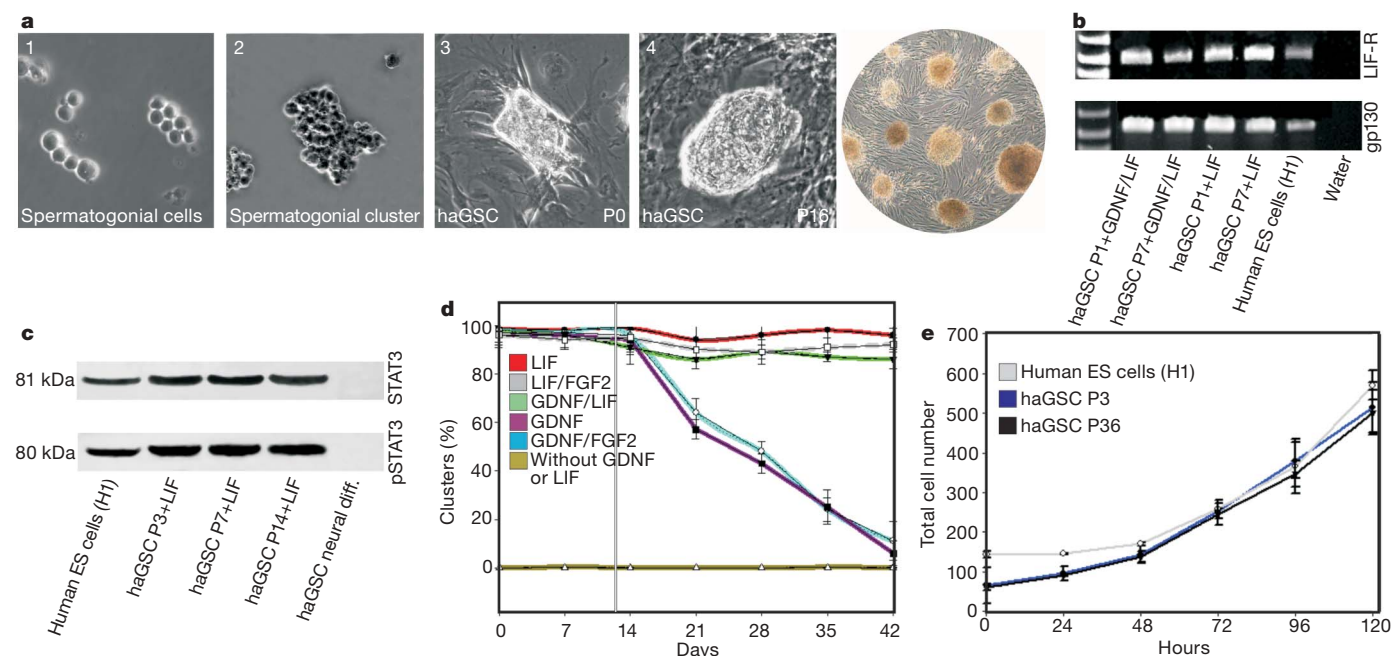


Figure 2 | Generation of human adult GSCs from spermatogonial cells.

a, Panel 1 shows colonies of spermatogonial cells on laminin (Lam_B cells). Panel 2 shows a proliferating spermatogonial cell colony. Panel 3 shows an early human adult GSC cluster (haGSC; P0). A human adult GSC cluster at higher passage (P16; panel 4) and an overview of typical human adult GSC clusters (far right panel) is also shown. **b**, RT-PCR of LIF receptor complex in human adult GSCs cultured under GDNF/LIF or LIF alone from lower (P1) and higher (P7) passage in comparison to FGF2-cultivated H1 cells. A

1-kilobase DNA ladder is shown. **c**, Western blot analysis of human adult GSCs under LIF from passage 3, 7 and 14, H1 cells and neural differentiated spermatogonial cells for 14 days under different growth factors over a further cultivation time period of 42 days expressed. **d**, Percentage of formed clusters after plating spermatogonial cells for 14 days under different growth factors over a further cultivation time period of 42 days expressed. **e**, Comparison of the doubling times of human ES cells (H1) and human adult GSCs from passage 3 and 36 over a 120-h period. Error bars in **d** and **e** show standard deviations ($n = 3$).

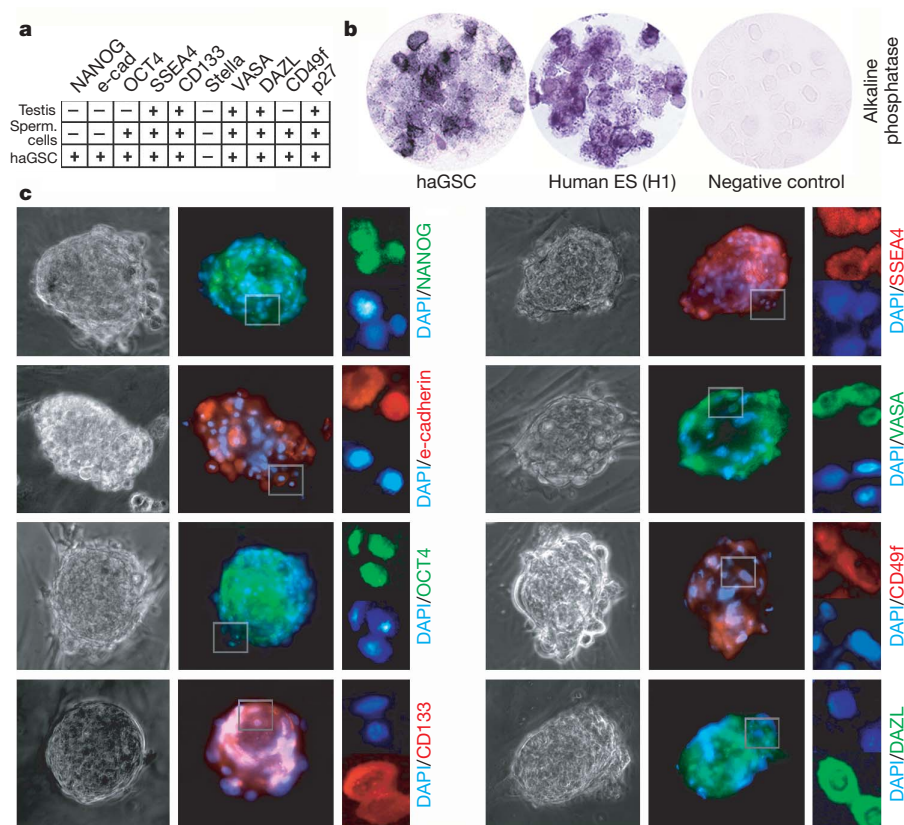


Figure 3 | Characterization of the human adult GSCs in comparison to normal testis tissue.

a, Comparison of stem cell, testis and spermatogonial markers in normal human testis, spermatogonial cells and human adult GSCs. **b**, Alkaline phosphatase staining of human adult GSC and human ES dissociated clusters as cytopins and HEK293 cells as negative control. Alkaline phosphatase was strongly positive in human adult GSC and human ES cell clusters. **c**, Immunostaining of human adult GSC clusters with NANOG, e-cadherin, OCT4, CD133, SSEA4, VASA, CD49f and DAZL. Columns from left to right: left, bright field; middle, double staining of stem cell marker with DAPI with areas of magnification; right, magnification of boxed area showing nuclear or cytoplasmic staining.

GDNF alone and GDNF with FGF2 clusters were formed, but the proliferation rate declined after 14 days. The combination of LIF and FGF2 did not improve human adult GSC culture.

To prove the dependence of our stem cells on LIF supplementation, LIF and gp130 receptors were detected in human adult GSCs with polymerase chain reaction with reverse transcription (RT-PCR; Fig. 2b). Furthermore, western blot analysis of STAT3 and phosphorylated STAT3 showed that this pathway is activated in human adult GSCs. These observations indicate that LIF has a role in human adult GSC propagation.

To test the expansion rate we compared the doubling times of different passages of human adult GSCs to those of human ES (H1) cells, counting the total cell numbers at 24-h intervals for approximately 120 h. The doubling times of human adult GSCs were very similar to those of human ES cells across lower and higher passages (Fig. 2e).

Human adult GSC characterization

For a more detailed examination of the human adult GSC colonies we performed immunohistochemical analysis and compared them to selected spermatogonial cells and normal adult human testis (Figs 1, 3a, c and Supplementary Figs 2 and 3). In normal adult human testis tissue, spermatogonia were positive for CD133, SSEA4, VASA, DAZL, TSPYL2 and CD49f, but negative for NANOG, e-cadherin, OCT4 and Stella (Supplementary Fig. 2). The same staining pattern was shown by the purified human spermatogonial cells, except that the transcription factor OCT4 started to be expressed (Figs 1 and 3a). Only morphologically typical Sertoli cells in the human testis stained positive for p27 (Supplementary Fig. 2). Alkaline phosphatase was highly expressed in cytopins of undifferentiated human adult GSCs and human ES cells, in contrast to differentiated HEK293 control cells (Fig. 3b). In contrast, the generated human adult GSC colonies were strongly positive for NANOG, e-cadherin, OCT4, CD133 and SSEA4 (ES cell markers) (Fig. 3a, c), but negative for p27 (a Sertoli cell marker) and Stella (a PGC marker) (Fig. 3a and Supplementary

Fig. 3i–j, panel 1). The human adult GSCs were also strongly positive for VASA, CD49f and DAZL (Fig. 3a, c).

Fluorescence-activated cell sorting (FACS) analysis provided quantitative information on the proportion of human adult GSC cells under LIF-supplemented culture conditions expressing particular surface markers in comparison to H1 (Supplementary Fig. 4a). These human ES cells typically expressed CD90, CD133, NANOG, e-cadherin, OCT4, SSEA4, SSEA3, TRA 1–60, TRA 1–81 and phosphorylated STAT3, and were negative for CD34, CD45, CD105 and CD117. In comparison, the expression profile of human adult GSCs was very similar. The only difference observed was the expression of CD117 (c-Kit).

RT-PCR analysis was carried out with complementary DNA from adult human testis, spermatogonial cells, human ES (H1) and human adult GSCs (under two different medium conditions) from lower and higher passages (Supplementary Fig. 4b). In adult human testis and spermatogonial cells, stem cell and spermatogonial cell markers (OCT4, n-cadherin, STAT3, SOX2, AFP, Stella, DAZL, VASA) were expressed, but not the markers NANOG, e-cadherin and GDF3. In contrast, human adult GSCs expressed all of these markers. Under LIF supplementation, the cultured human adult GSCs started to express the ES cell factors NANOG, e-cadherin and GDF3 as well, markers which were not expressed in the adult human testis and spermatogonial cells.

Western blot analysis also showed that NANOG, e-cadherin, OCT4 and SSEA4 were expressed by human adult GSCs under LIF supplementation (Supplementary Fig. 4c). These results indicate that human adult GSCs respond to culture conditions and acquire more human ES cell properties.

The expression of spermatogonial and human ES cell markers is maintained even in higher passages. This indicates that the cultured human adult GSCs have a stable, stem-cell-like phenotype.

In a next step we investigated and compared the global gene expression patterns of spermatogonial cells, human ES cells and human adult GSCs by microarray analysis (Fig. 4a). After principal

component analysis (PCA) of germ- and stem-cell-specific transcripts we found very similar gene expression profiles of human adult GSCs and human ES cells with minor differences. In contrast, the expression profile of spermatogonial cells changed during development to human adult GSCs. The gene products in which expression was found to differ significantly between human ES cells and human adult GSCs are involved in tissue, organ and embryonic development, cell cycle, cellular assembly and organization, and cellular growth and proliferation (Supplementary Fig. 5a). Expression profiles of human ES and germ cell markers in the three different cell types are shown in Supplementary Table 2.

Among the genes most significantly upregulated in human adult GSCs in comparison to human ES cells are members of the WNT- β -catenin and TGF- β signalling pathways (Supplementary Fig. 5b).

Figure 4b shows selected examples of stem cell and testis marker expression profiles from the molecular analysis comparing human adult GSCs, human ES cells and spermatogonial cells. Expression analysis of human adult GSCs compared with human ES cells and spermatogonial cells revealed high levels of *STAT3*, *CD9*, *KLF4*, *OTEX* and *VASA* in both human adult GSCs and human ES cells. *POU6F1* and *DAZL* expression was stronger in human adult GSCs. *OCT4*, e-cadherin, *NANOG*, *SOX2* and *GDF3* expression was stronger in human ES cells. Minimal expression of the germ cell marker *Stella* was seen in all cell types. *SPAG9* and *TSPYL2* were more highly expressed in spermatogonial cells, which clearly indicates their germline origin.

The real-time PCR analysis demonstrated that human ES cells and human adult GSCs both express a similar transcription profile of markers used to characterize human ES cells (Fig. 4c). The alignment of genes showed few differences and the human adult GSCs differed from human ES cells only in expression potency. Taken together, the examinations of transcription in human ES and human adult GSCs provide evidence of a network of genes characteristic of pluripotent stem cells.

To determine whether the maintenance of genomic imprinting was compromised we assessed the methylation pattern of the imprinted genes *H19* and *IGF2R* with DNA isolated from human ES cells (H1,

P48), spermatogonial cells (P0) and human adult GSCs from four different normal patients obtained from two passages (LP, P1–P3; HP, P7–P36). As shown in Supplementary Fig. 5c, spermatogonial cells showed a typical methylation pattern known for male germ cells, with hypermethylation of the maternal differentially methylated region (DMR) located at the 5' untranslated region (UTR) of the *H19* gene and no methylation of the paternal DMR imprinting control region surrounding the transcriptional start site (TSS) of the *IGF2R* gene. The same methylation was observed for undifferentiated H1 cells which have a normal XY karyotype.

In general, both spermatogonial and H1 cells were hypermethylated (>90%) at the 5' UTR of the *H19* gene and showed about 75% methylation in human adult GSCs. The region surrounding the TSS of the *IGF2R* gene was unmethylated (<5%) in all samples studied. We did not find differential methylation among the different human adult GSC groups in any of the analysed regions when samples were grouped according to cluster or when samples of lower and higher passages of human adult GSCs were compared (data not shown). These observations indicate that human adult GSCs change their methylation pattern in the DMR of imprinting sites and display a pattern similar to that of mouse spermatogonial stem cells as shown previously⁵.

In addition, when analysing other sites of expected differential methylation, H1 cells showed higher methylation (approximately 75%) than spermatogonial cells and human adult GSCs (approximately 50%) in the region surrounding the TSS of the *H19* gene. In turn, the imprinting controlling region (ICR) within intron 2 of the *IGF2R* gene was differentially methylated in spermatogonial cells, H1 and human adult GSCs (approximately 50%, 100% and 75%, respectively).

Next we investigated DNA methylation levels on the putative human *OCT4* (also called *POU5F1*) and *NANOG* promoters. As shown in Fig. 4d, when DNA methylation results were grouped by cell type we found regions of differential methylation. The region located at the 5' UTR of the *OCT4* and *NANOG* genes showed cell-type-specific DNA methylation of spermatogonial cells, H1 cells

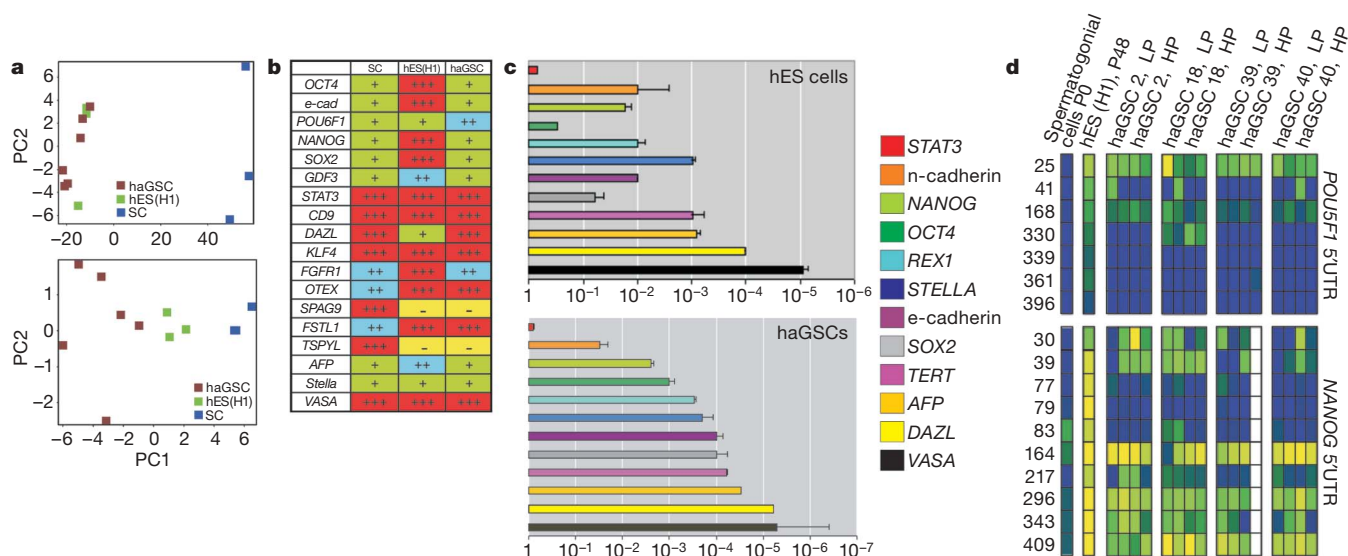


Figure 4 | Molecular profiling and epigenetics of human adult GSCs. **a**, PCA diagrams comparing spermatogonial cells (SC), human adult GSCs and human ES (hES) populations and all biological replicates (upper panel). Expression profile of germ- and stem-cell-specific transcripts revealed a similar expression pattern of human ES cells (green) and human adult GSCs (brown) (lower panel). Human spermatogonial cells (blue) differed significantly from the other samples. PC1, first principal component; PC2, second principal component. For all groups $n = 3$. **b**, Stem cell and testis expression profiles in all three cell types. The expression rate of genes is shown in different signs and colours: -/yellow, not expressed; +/light green, low; ++/blue, moderate; +++/red, high. **c**, Comparison of gene expression between human ES cells and human adult GSCs by real-time PCR. Amplified genes are shown in different colours. The messenger RNA levels were normalized to *GAPDH* and error bars show standard deviations ($n = 3$). **d**, Average methylation of *POU5F1* and *NANOG* promoter genes with sample grouping by cell type. Colour codes indicate 0% (yellow) over 50% (green) to 100% (dark blue) methylation. The y axis shows amplificate location at individual CpG positions. The x axis shows samples that are grouped by cell type and passage (LP, passage 3; HP, passage 7).

low; ++/blue, moderate; +++/red, high. **c**, Comparison of gene expression between human ES cells and human adult GSCs by real-time PCR. Amplified genes are shown in different colours. The messenger RNA levels were normalized to *GAPDH* and error bars show standard deviations ($n = 3$). **d**, Average methylation of *POU5F1* and *NANOG* promoter genes with sample grouping by cell type. Colour codes indicate 0% (yellow) over 50% (green) to 100% (dark blue) methylation. The y axis shows amplificate location at individual CpG positions. The x axis shows samples that are grouped by cell type and passage (LP, passage 3; HP, passage 7).

and human adult GSCs. Whereas spermatogonial cells were hypermethylated (100%) at the 5' UTR of the *OCT4* gene, H1 cells (60%) and human adult GSCs (50%) were less methylated at CpG genomic position 25–168, but were hypermethylated (100%) between positions 330 and 396. In addition, spermatogonial cells were again hypermethylated (approximately 90%) at the 5' UTR of the *NANOG* gene, whereas H1 cells displayed low methylation (approximately 20%) and human adult GSCs also showed clearly lower (approximately 40%) methylation than spermatogonial cells. These results indicate that the epigenetic state of the *OCT4* and *NANOG* promoter genes was reprogrammed to a more human ES-like state when human adult GSCs were generated from spermatogonial cells.

Human adult GSCs were also examined for their pluripotency in forming teratomas *in vivo* after injection in immunodeficient mice. The human adult GSC cells from eight different patients were injected ($n = 4$ sets of cells per patient) and gave rise to typical teratoma structures in nude mice by 6 to 8 weeks after the transplantation (Fig. 5, Supplementary Fig. 6 and Supplementary Table 3). The tumours contained foci with derivatives of endodermal, mesodermal and ectodermal embryonic germ layers: stratified cell epithelium (Fig. 5, panel 1 and Supplementary Fig. 6a, panel 1), neuronal cells (Supplementary Fig. 6a, panel 1), cartilage, muscle (Fig. 5, panel 2 and Supplementary Fig. 6a, panel 2), glandular structures (Fig. 5, panel 3 and Supplementary Fig. 6a, panel 3) and endodermal high prismatic epithelium (Fig. 5, panel 3). Immunostaining was carried out with the germ layer markers cytokeratin for epithelial tissue in endoderm, TUJ-1 for neuroectodermal cells and α -fetoprotein (AFP) for endoderm (Supplementary Fig. 6a, panels 1 and 3). Alcian blue stained hyaline cartilage (Supplementary Fig. 6a, panel 2) and epithelial glands (Supplementary Fig. 6a, panel 3).

In general, the generation of human adult GSCs from spermatogonial cells was reproducible: similar cell numbers were obtained from biopsies obtained from the 22 male patients (10 normal, 10 azoospermia, 2 sex reassignment surgeries) aged 17 to 81 years (see Supplementary Table 1 for an overview of number of patients and human adult GSC cell cultures used for experiments). After selection the obtained yields of spermatogonial cells were age-dependent. Lesser amounts were generated from older people and patients with azoospermia. However, the doubling times of human adult GSC colonies were similar in all groups (data not shown). Human adult GSCs were successfully passaged for continuous undifferentiated proliferation in basic medium with LIF for up to 16 months and over more than 40 passages. We did not observe a decline in the ability to form colony units between passages 3 and 36 (Fig. 2e). The colonies had a doubling time of 48 h. The undifferentiated human adult GSCs could be cryopreserved and thawed with no loss of proliferation or differentiation capacity. No replicative crisis was observed in any of the 22 different cell lines. Cytogenetic analysis showed that human adult GSCs had a normal karyotype (46 chromosomes, XY) in all examined metaphase spreads (Supplementary Fig. 7).

To determine whether human adult GSCs can differentiate *in vitro*, we applied commonly used methods designed to induce differentiation of human ES cells into various cell lineages. On the whole, differentiated human adult GSCs displayed all morphological characteristics of the expected myogenic, osteogenic, pancreatic and neural lineages (Supplementary Fig. 8).

To analyse the differentiation capacity of human ES (H1) and human adult GSCs, the cells were differentiated into all specific lineages (Supplementary Fig. 9). The results were comparable, diverging only in the pancreatic lineage, where we generated more insulin-, glucagon- and c-peptide-positive cells from human adult GSCs.

To exclude possible contaminations during cell culturing, DNA microsatellite markers were analysed in DNA preparations of all established cell cultures and compared with a marker profile of H1 embryonic stem cells. All cell cultures were analysed after completion of all other experiments and showed independent genetic origins (Supplementary Table 5), that is, each cell line is unique.

Moreover, we never observed more than two marker alleles, thus excluding relevant cell or DNA contamination for all cell lines.

Discussion

We have developed a culture method for establishing human adult GSCs from testicular biopsies. We can isolate and cultivate over a long period of time (over 40 passages and more) highly proliferative, stable human adult GSC cell cultures derived from pure enriched spermatogonial cells of different adult human testes under LIF

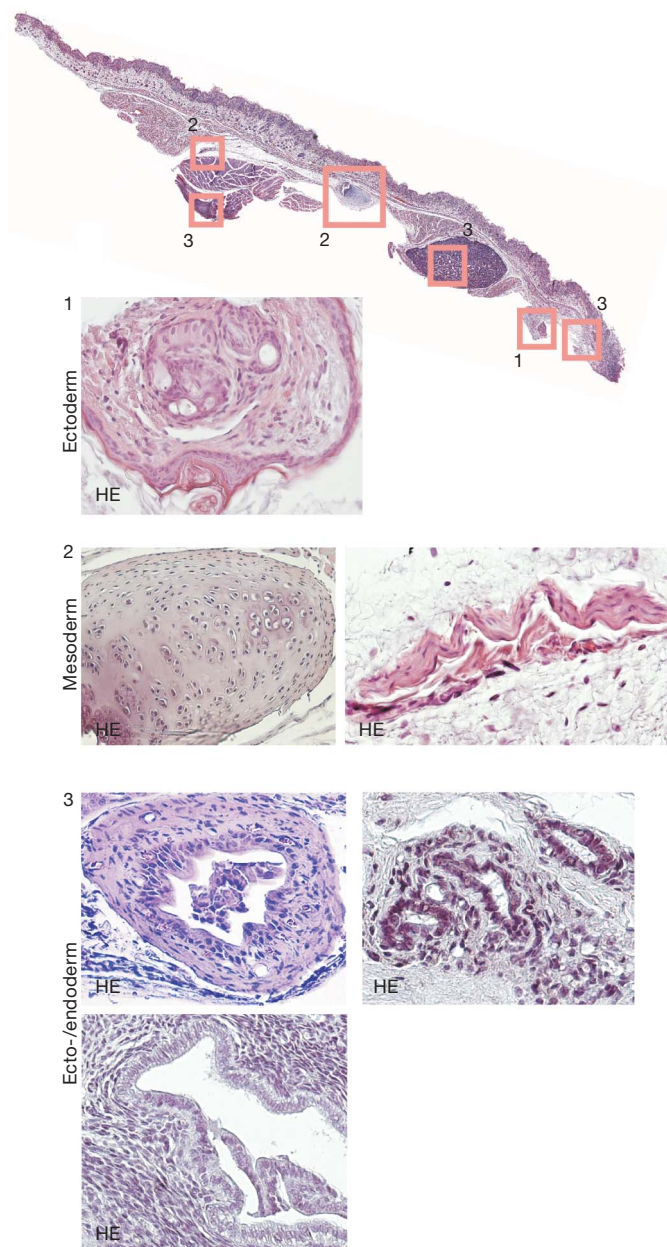


Figure 5 | Human adult GSC-derived human teratoma formation after injection into an immunodeficient mouse. Complete section (low-resolution image; mosaic scan with $\times 40$ original magnification) of the gross anatomy of the subcutaneous tumour formation is shown at the top; higher-magnification images ($\times 63$ original magnification) of boxed areas are shown below. Panel 1 shows formation of ectodermal structures (stratified keratinizing epithelium developed). Panel 2 shows mesodermal structures from human adult GSCs with cartilage and smooth muscle. Panel 3 shows development of endodermal structures with pseudostratified columnar epithelium (top left), glandular structures (top right) and gut-like simple columnar epithelium (bottom left). HE, haematoxylin and eosin stain.

supplementation, a growth factor also important for the derivation of human embryonic germ cells from PGCs^{13,14}.

The spermatogonial stem cells were isolated with CD49f MACS separation, followed by sequential matrix selection with collagen and laminin. In the absence of highly selective antibodies for human spermatogonial stem cells^{15–18}, matrix selection is an important tool to gain a pure spermatogonial stem cell population¹⁹. In the future, mass spectroscopy screening of proteins expressed on the cell surface will be used to generate highly specific antibodies against human spermatogonial stem cells that can be used for optimal pre-selection. Molecular analysis of the human adult GSCs showed that these cells changed their properties, losing characteristics of spermatogonial cells and acquiring characteristics and expression profiles similar to those of human ES cells. Human adult GSCs express genes that have recently been shown to be necessary to reprogram adult somatic cells to an ES-like state^{20–24}. Like germ stem cells from neonatal and adult mice, human adult GSCs are highly versatile and can give rise to different cell types of all germ layers. Because we cannot inject the human adult GSCs into blastocysts to interrogate their pluripotency, *in vitro* differentiation capacity and the teratoma assay were conducted for functional proof of the pluripotency of human adult GSCs.

There are considerable differences between mouse and human ES cells²⁵, and presumably between adult mouse and human GSCs as well (refs 5, 6, 10 and this study). The precise molecular and cellular mechanisms that specifically regulate the proliferation and pluripotency of human adult GSCs remain unclear. The generation of human adult GSCs from VASA⁺ vimentin[−] spermatogonial cells was highly dependent on culturing in the presence of LIF. In the future, more research will need to be conducted on culture conditions with different combinations of growth factors, the function of protein interactions and the need for feeder layers. The availability of histocompatible, pluripotent cells for autotransplantation from simple biopsies would circumvent immunological problems associated with regular human ES-cell-based technology. Further investigations will be needed to analyse whether these cells are different from other recently reported mouse and human pluripotent cells obtained by reprogramming in terms of morphology, marker expression, genetic imprinting, capacity for differentiation and genetic stability. Such studies will provide important information for potential clinical applications.

METHODS SUMMARY

Isolation of spermatogonial cells and generation of human adult GSCs. In total 22 different human testicular parenchymas prepared from biopsies or orchiectomy specimens were used for the study. Any abnormality was excluded by pathological examination except for azoospermia with residual spermatogenesis. After mechanical and enzymatic digestion, single cells were plated in knock-out medium with 4 ng ml^{−1} GDNF (Sigma) and incubated at 32.5 °C, 5% CO₂ for 96 h. For selection of spermatogonial cells, a combination of MACS CD49f pre-selection in combination with a matrix purification protocol of LamB cells¹⁹ was used. The LamB cells were harvested and cultured in basic medium with LIF on 0.1% gelatine-coated plates. Under these conditions the spermatogonial cells first proliferated and aggregated, and then pluripotent human adult GSC clusters were formed in 21–28 days. Generally the cells were passaged mechanically every 14 days. Only cell cultures with a normal karyotype (46,XY in all of the examined metaphase spreads) in low (P3) and high (P36) passages were used.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 1 February; accepted 18 September 2008.

Published online 8 October 2008.

- de Rooij, D. G. Rapid expansion of the spermatogonial stem cell tool box. *Proc. Natl Acad. Sci. USA* **103**, 7939–7940 (2006).
- Brinster, R. L. & Avarbock, M. R. Germline transmission of donor haplotype following spermatogonial transplantation. *Proc. Natl Acad. Sci. USA* **91**, 11303–11307 (1994).

- Kubota, H. & Brinster, R. L. Technology insight: *In vitro* culture of spermatogonial stem cells and their potential therapeutic uses. *Nature Clin. Pract. Endocrinol. Metab.* **2**, 99–108 (2006).
- Matsui, Y., Zsebo, K. & Hogan, B. L. Derivation of pluripotent embryonic stem cells from murine primordial germ cells in culture. *Cell* **70**, 841–847 (1992).
- Kanatsu-Shinohara, M. *et al.* Generation of pluripotent stem cells from neonatal mouse testis. *Cell* **119**, 1001–1012 (2004).
- Guan, K. *et al.* Pluripotency of spermatogonial stem cells from adult mouse testis. *Nature* **440**, 1199–1203 (2006).
- Stevens, L. C. Spontaneous and experimentally induced testicular teratomas in mice. *Cell Differ.* **15**, 69–74 (1984).
- Resnick, J. L., Bixler, L. S., Cheng, L. & Donovan, P. J. Long-term proliferation of mouse primordial germ cells in culture. *Nature* **359**, 550–551 (1992).
- Turnpenny, L. *et al.* Evaluating human embryonic germ cells: concord and conflict as pluripotent stem cells. *Stem Cells* **24**, 212–220 (2006).
- Seandel, M. *et al.* Generation of functional multipotent adult stem cells from GPR125+ germline progenitors. *Nature* **449**, 346–350 (2007).
- Meng, X. *et al.* Regulation of cell fate decision of undifferentiated spermatogonia by GDNF. *Science* **287**, 1489–1493 (2000).
- Kanatsu-Shinohara, M. *et al.* Leukemia inhibitory factor enhances formation of germ cell colonies in neonatal mouse testis culture. *Biol. Reprod.* **76**, 55–62 (2007).
- Shamblott, M. J. *et al.* Derivation of pluripotent stem cells from cultured human primordial germ cells. *Proc. Natl Acad. Sci. USA* **95**, 13726–13731 (1998).
- Shamblott, M. J. *et al.* Human embryonic germ cell derivatives express a broad range of developmentally distinct markers and proliferate extensively *in vitro*. *Proc. Natl Acad. Sci. USA* **98**, 113–118 (2001).
- Costoya, J. A. *et al.* Essential role of Plzf in maintenance of spermatogonial stem cells. *Nature Genet.* **36**, 653–659 (2004).
- Shinohara, T., Avarbock, M. R. & Brinster, R. L. β_1 - and α_6 -integrin are surface markers on mouse spermatogonial stem cells. *Proc. Natl Acad. Sci. USA* **96**, 5504–5509 (1999).
- Kubota, H., Avarbock, M. R. & Brinster, R. L. Culture conditions and single growth factors affect fate determination of mouse spermatogonial stem cells. *Biol. Reprod.* **71**, 722–731 (2004).
- Stukenborg, J. B. *et al.* Co-culture of spermatogonia with somatic cells in a novel three-dimensional soft-agar-culture-system. *J. Androl.* **29**, 312–329 (2007).
- Hamra, F. K. *et al.* Defining the spermatogonial stem cell. *Dev. Biol.* **269**, 393–410 (2004).
- Wernig, M. *et al.* *In vitro* reprogramming of fibroblasts into a pluripotent ES-cell-like state. *Nature* **448**, 318–324 (2007).
- Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**, 1917–1920 (2007).
- Park, I. H. *et al.* Reprogramming of human somatic cells to pluripotency with defined factors. *Nature* **451**, 141–146 (2008).
- Meissner, A., Wernig, M. & Jaenisch, R. Direct reprogramming of genetically unmodified fibroblasts into pluripotent stem cells. *Nature Biotechnol.* **25**, 1177–1181 (2007).
- Okita, K., Ichisaka, T. & Yamanaka, S. Generation of germline-competent induced pluripotent stem cells. *Nature* **448**, 313–317 (2007).
- Giniis, I. *et al.* Differences between human and mouse embryonic stem cells. *Dev. Biol.* **269**, 360–380 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank S. Singer, U. Mau-Holzmann and O. Rieß for karyotyping the human adult GSCs obtained from the patients' biopsies; M. Schenk for providing support with ELISA; D. Blaurock for proofreading the manuscript; K. Kohler for discussion and encouragement; and P. Bauer for conducting the microsatellite analysis. Financial support for this research was provided by the Department of Urology, Institute of Anatomy and ZRM (University Clinic Tübingen).

Author Contributions S.C. conducted all the major experiments and wrote the manuscript; M.R. and Jö.H. were responsible for the clinical logistics, organization of the human testis biopsies and writing the ethical proposal; L.J. supervised the experimental procedures conducted in Tübingen; M.R., M.M. and Jö.H. provided mRNA and cDNA from human ES cells (H1), determined the doubling time of H1 cells and were involved in the differentiation experiments for comparing H1 with human adult GSCs; T.W. conducted part of the PCR experiments; W.A. was responsible for the real-time PCR experiments, M.B. was responsible for microarray profiling, H.-J.B. was co-responsible for FACS analysis; U.M. and H.-J.W. conducted the electron microscopy study; A.M. performed confocal microscopy; S.M. provided human ES cells; K.-D.S. and A.S. supported and advised the clinical cooperation; and T.S. initiated and supervised the entire project, conducted parts of the experiments and wrote the manuscript. A.S. and T.S. are co-senior authors.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to T.S. (Thomas.Skutella@regmed.uni-tuebingen.de).

METHODS

Cell culture of human adult GSCs. The experiments with human material were approved by the local ethics council (University Clinic Tübingen). Informed consent was obtained from all the human subjects.

The obtained human testis tissues were mechanically disrupted and enzymatically dissociated with 0.5 g ml⁻¹ collagenase type VI (Sigma) and 0.25 g ml⁻¹ dispase II (Roche) in HBSS buffer with Ca²⁺ and Mg²⁺ (PAA) for 30 min at 37 °C. Then the digest was pelleted at 1,000 r.p.m., washed twice with knockout culture medium (Gibco) with 20% ES cell qualified FBS and 1% L-glutamine and filtered through a 40 µm mesh to obtain a single cell suspension. Cells were plated into 10 cm² culture dishes at 2 × 10⁶ cells per cm² containing medium with 4 ng ml⁻¹ GDNF (Sigma) and incubated at 32.5 °C, 5% CO₂ for 96 h. Medium was removed after 96 h and testis cell cultures were gently washed with 4 ml of DMEM high glucose (PAA) and once with 4 ml of PBS. Bound germ cells were harvested from monolayers of adherent somatic cells by repeated pipetting with 4 ml of DMEM. The pooled suspension was pelleted at 1,000 r.p.m., suspended in 10 ml DMEM and filtered through a 40-µm mesh. For further purification MACS separation (Miltenyi) with biotinylated CD49f (α₆ integrin; BioLegend) and anti-biotin beads was applied^{5,15,16}. After MACS separation, cells from five 10-cm dishes were transferred to a 10-cm plastic dish coated with collagen I (5 µg cm⁻², Becton Dickinson) and incubated at 32.5 °C for 4 h in basic medium (DMEM high glucose, 15% FCS (Biochrom), 1% non-essential amino acids (NEAA), 1% L-glutamine and 0.05 mM β-mercaptoethanol (Gibco) with 10³ units ml⁻¹ leukaemia inhibitory factor (LIF, human, Chemicon))⁵. Cells that did not bind to collagen I dishes (Col_{NB} cells) were harvested and pelleted at 1,000 r.p.m. The Col_{NB} cells were suspended in basic medium and plated at 0.5–1 × 10⁶ cells per ml per well in 12-well plates pre-coated with laminin (4.4 µg cm⁻², Sigma). The plated Col_{NB} cells were incubated for 45 min at 32.5 °C and unbound cells (Col_{NB}/Lam_{NB} cells) were removed from bound cells (Lam_B cells) by pipetting and discarded. The Lam_B cells were rinsed twice with 1 ml DMEM. The Lam_B cells then were harvested by gentle pipetting and cultured in basic medium on 0.1% gelatine-coated 48-well plates. Under these conditions the spermatogonial cells proliferated first and aggregated, and then human adult GSC clusters were formed. We only used a few (500) purely isolated spermatogonial cells to generate human adult GSC cultures from the patients in 48-well plates. Generally the cells were passaged mechanically every 14 days. Clusters were manually isolated and cut into pieces and plated on gelatine-coated dishes. Only cell cultures with a normal karyotype (46,XY in all of the examined metaphase spreads) in low (P3) and high (P36) passages were used.

For purposes of comparison all the experiments were also performed in parallel with the human ES cell line H1 in the laboratory of J. Hescheler (Cologne, Germany). The H1 cells were cultured in knockout DMEM (Invitrogen), 20% knockout serum replacer (Invitrogen), 1 mM L-glutamine (Invitrogen), 1% NEAA (Invitrogen), 0.1 mM β-mercaptoethanol (Gibco), 1% penicillin/streptomycin (Gibco) and 4 ng ml⁻¹ FGF2 (Peprotech) on CF1-inactivated feeders on 0.1% gelatine-coated dishes and passaged mechanically every 6–7 days. Colonies were cut manually into pieces, counted and distributed on new feeder-coated dishes (~1.34 × 10⁴ cells cm⁻²)²⁶.

Testing of different culture conditions. Different medium combinations were tested with spermatogonial cells cultured on 0.1% gelatine-coated dishes or on different coatings. To this end spermatogonial cells were plated for 14 days and the formed clusters observed for another 42 days under 8 different conditions. (1) Basic medium directly from the beginning of culture with LIF (10³ units ml⁻¹); (2) condition 1 but with FGF2 (4 ng ml⁻¹); (3) knockout medium with GDNF (4 ng ml⁻¹) for the first 4 days followed by basic medium with LIF; (4) knockout medium with GDNF; (5) condition 4 but with FGF2 (4 ng ml⁻¹); (6) knockout medium or basic medium without GDNF, FGF2 or LIF; (7) condition 1 or 3 on a combination of laminin, fibronectin and poly-L-ornithine; (8) condition 1 or 3 on matrigel.

Quantification of the doubling time of human adult GSCs and human ES cells. For this experiment 3 × 10⁵ cells per plate from three different passages of human adult GSCs and human ES cells were seeded in 6-well plates. Clusters were cut from human adult GSCs and human ES cultures, digested, counted and replated every 24 h over a time period of 120 h.

Differentiation of human adult GSCs and human ES cells. The isolated cells were transposed into specific established culture media to initiate differentiation into all three germ layers. We used specific protocols for myogenic^{27,28}, osteogenic^{29,30}, pancreatic^{31–33} and neural^{34,35} differentiation of human ES cells. All the differentiation protocols were repeated at least three times under each set of conditions using mechanically isolated clusters from human adult GSCs and human ES (H1) cells.

Alkaline phosphatase staining of cytopins of human adult GSCs and human ES cells. For staining, clusters were mechanically isolated and dissociated with trypsin EDTA for 5 min at 37 °C. After blocking with 10% FCS and washing with PBS, 1 × 10⁶ cells were resuspended in 0.5 ml PBS and centrifuged as cytopins for 4 min at 1,100 r.p.m. Then the spins were fixed with 2% paraformaldehyde for 5 min and stained with NBT/BCIP. HEK293 cells were used as controls.

Immunohistochemistry. To characterize human spermatogonia, human adult GSCs and the differentiation into derivatives of the three primary germ layers, we examined the expression of a panel of cell-specific proteins for spermatogonial cells, human ES cells and markers of myogenic, osteogenic, pancreatic and neural differentiated cells.

Antibodies and staining. The following primary antibodies were used: mouse monoclonal biotinylated anti-CD49f (BioLegend), rabbit polyclonal anti-OCT4 (Abcam), mouse monoclonal anti-SSEA4 (Chemicon), rabbit polyclonal anti-NANOG (Biozol), mouse monoclonal anti-e-cadherin (R&D Systems), goat polyclonal anti-human VASA (R&D Systems), rat monoclonal anti-Stella (R&D Systems), rabbit polyclonal anti-p27 (Abcam), mouse monoclonal anti-SOX17 (R&D Systems), goat polyclonal anti-FOXA2 (R&D Systems), rat monoclonal anti-CXCR4 (BD Pharmingen), mouse monoclonal anti-GSC (Abnova), mouse monoclonal anti-human c-peptide (BioVendor), rabbit polyclonal anti-insulin (Santa Cruz), rabbit polyclonal anti-human glucagon (Dako), mouse monoclonal anti-α-actinin (Sigma), rabbit polyclonal anti-smooth muscle actin (Spring Bioscience), rabbit polyclonal anti-glia fibrillary acidic protein (GFAP) (Dako), mouse monoclonal anti-β-tubulin-III (TUJ-1) (Convince), mouse monoclonal anti-neurofilament (Dako), mouse monoclonal anti-glutamate (Sigma), rabbit polyclonal anti-GABA (Sigma), rabbit polyclonal anti-cytokeratin (Dako), and mouse monoclonal anti-α-fetoprotein (Chemicon). Alexa Fluor-488-conjugated goat anti-mouse IgG, Alexa Fluor-488-conjugated goat anti-rabbit IgG, Alexa Fluor-488-conjugated goat anti-rat IgG (Molecular Probes), Cy3-conjugated goat anti-mouse IgG, Cy-conjugated goat anti-rabbit IgG and Cy-conjugated goat anti-rat IgG (Dianova) were used as secondary antibodies with co-staining with DAPI. In addition biotinylated swine anti-rabbit, biotinylated rabbit anti-mouse and biotinylated goat anti-rat (Dako) with ABC complex (streptavidin/horseradish peroxidase) and DAB staining with haemalaun or DAPI as co-staining were used. For negative controls, isotype mouse, goat, rat or rabbit IgGs were used. Alkaline phosphatase staining was carried out using NBT/BCIP substrate (Roche).

Western blot analysis. Clusters of human adult GSCs, H1, HEK293 and human adult GSC neural differentiated cells were lysed in RIPA buffer and sonicated. The probes were denatured in sample buffer, analysed in a 12% SDS-PAGE and after blotting stained with the specific antibodies for NANOG, OCT4, SSEA4, e-cadherin, STAT3 (Santa Cruz) and pSTAT3 (Santa Cruz). For negative control HEK293 or blocking peptides for STAT3 and pSTAT3 were used (data not shown).

RT-PCR analysis. Tissues and cells were homogenized in lysis buffer and total RNA was prepared using the RNeasy Mini Kit (Qiagen). cDNA for human ES cells was obtained from the stem cell line H1. mRNA was reverse transcribed using Oligo (dT)23 primer (Sigma) and SuperscriptII-transcriptase (Invitrogen). cDNA was amplified using the primers and conditions shown in Supplementary Table 1.

Real-time PCR analysis. Steady state mRNA levels were enumerated by quantitative RT-PCR (qRT-PCR, LightCycler, Roche) as described³⁶. The qRT-PCR was performed as touch-down PCR in 35 cycles³⁷. Quantification of *GAPDH* and serial dilutions of recombinant standard DNAs served as controls in each PCR. Transcript amounts are presented as copy numbers normalized to *GAPDH* and the recombinant standards. Analysis of melting curves confirmed product quality after each PCR.

Flow cytometric analysis. The following unconjugated and conjugated antibodies were used for FACS analysis: mouse monoclonal anti-CD34 (BD Pharmingen), anti-CD44-FITC (Becton Dickinson), mouse monoclonal anti-CD45 (BD Pharmingen), anti-CD90-APC (BD-Pharmingen), mouse monoclonal anti-CD105 (Ansell), anti-CD117-FITC (Chemicon), mouse monoclonal anti-CD133 (University Clinic of Tübingen), rat monoclonal anti-SSEA3 (R&D Systems), anti-SSEA4-APC (BD Pharmingen), mouse monoclonal TRA 1–60, mouse monoclonal TRA 1–81 (Santa Cruz), e-cadherin (R&D Systems) and the intracellular antibodies for OCT4, NANOG (Abcam) and pSTAT3 (Santa Cruz). For intracellular staining the cells were fixed with 2% paraformaldehyde and permeabilized with 0.1% Triton X-100 before labelling. At least 10,000 events were acquired on a FACSCanto II cytometer (Becton Dickinson) using the FACS express software for analysis.

Electron microscopy. After embedding in Epon (Fluka), semithin sections (1 µm) were prepared and from them ultrathin sections (100 nm). Semithin sections were stained with toluidine blue and embedded in epoxy resin for

ultrathin sectioning in electron microscopy. Digital micrographs were taken with a Zeiss Axioskop microscope.

Methylation assay. The methylation assay was conducted by Epigenomics (Berlin). Briefly, after quality control was performed for 18 DNA samples from spermatogonial cells (directly after matrix selection), human adult GSCs (from lower and higher passages), human ES cell line H1 and from human ES cells (obtained from S. Minger), the genomic DNA was treated with sodium bisulphite. PCR primers were designed for the regions of interest in the specified genes. Two regions of interest in the *IGF2R* and *H19* genes were studied. The first region in the *H19* gene is 550 bp upstream of the Vega H19-012 transcript within a CpG-rich region and at the 5' UTR of the annotated Ensembl transcript NR_002196.1 within a CpG-rich region. It contains the TSS of the reference sequence. The second region of interest in the *IGF2R* gene is 273 bp upstream of the annotated Ensembl transcript MPRI_HUMAN near a CpG island and within the imprinting controlling region 1 (ICR1) and a CpG island in intron 2. This region is known as imprinting controlling region 2 (ICR2). The following regions of interest in *POU5F1* (annotated *OCT4* orthologous human gene) and *NANOG* genes were studied: *POU5F1* gene (reference sequence: NM_002701) AMP1000122 located at the 5' UTR of the annotated Ensembl transcript POUF1_HUMAN (ENST00000259915), 150 bp upstream of the TSS. *NANOG* gene (reference sequence: NM_024865) AMP1000123 located at the 5' UTR of the annotated Ensembl transcript NANOG_HUMAN (ENST00000229307), 25 bp upstream of the TSS. The following bisulphite primers were used for PCR and for sequencing: *H19* (5' UTR) 5'-ATATTGAAGTTTTAGAG-TGTGATTT-3' and 5'-TTCCCCTTCTATCTCACCA-3'; *IGF2R* (TSS) 5'-TT-TTATTTTGTGGATTGTGTT-3' and 5'-AACCTCAATTTCCCCTCC-3'; *H19* (TSS) 5'-GGAGATAGTGGTTTGGGAG-3' and 5'-ACCCCATCTTCCC-TAAT-3'; *IGF2R* (intron 2) 5'-GGTGTAGGGGATTTAGGG-3' and 5'-AAAC-CTTTTCTACCTCCCTTT-3'; *POU5F1* 5'-ATGGTGTTTGTGGAAGGGG-AA-3' and 5'-TCCAAACAATAAAATATACAAAACCT-3'; *NANOG* 5'-TAA-TATGAGGTAATTAGTTTGTAGTTAGT-3' and 5'-TAATTTCAAACCTAAC-TTCAAATAAT-3'.

In addition, DNA from peripheral blood lymphocytes (sample name Pro) and artificially hyper- and hypomethylated DNA samples (sample names Up and Down, respectively) were included as controls. Bisulphite-converted DNA was amplified using these primers and the PCR products were directly sequenced. Methylation results were quantified using Epigenomics' proprietary software ESME, which allows the quantification of DNA methylation at single CpG positions. The statistical significance of the differential methylation observed between sample groups was tested using Wilcoxon's rank sum test.

Microarray analysis. Total RNA was isolated from $n = 3$ independent cell preparations of spermatogonial cells (directly after matrix selection), human adult GSCs (from lower and higher passages), human ES cell line H1 and from human ES cells (obtained from S. Minger) using the RNeasy Mini Kit (Qiagen) followed by an amplification step with the MessageAmp aRNA Kit (Ambion). Samples were analysed independently. Gene expression analyses were performed using the Human U133 + 2.0 Genome oligonucleotide array (Affymetrix) and all transcript intensities were GC-RMA normalized and analysed with a *t*-test using the ArrayAssist 4.0 software (Stratagene). Transcripts with more than a twofold estimated difference in expression were further considered for evaluation. For stringent multiple testing correction we used the Bonferroni-Holm procedure. In addition, we applied the web-based interactome network entry tool developed by Ingenuity Systems.

Insulin ELISA. Quantitative insulin determination by ELISA was performed as described previously³³. The supernatants were analysed with the insulin ELISA kit (YK060, Cosmo, Ltd) and the ultrasensitive c-peptide ELISA kit (Mercodia).

Transplantation procedures. For all implantation and transplantation experiments the mice were anaesthetized with an isoflurane inhalation system. The experimental protocols were approved by the local governmental council for animal care (Regierungspräsidium Tübingen) and were conducted according to the German law for the care and use of laboratory animals.

Transplantation of human adult GSCs in nude mice. Approximately 1×10^7 human adult GSCs were injected intramuscularly or subcutaneously in the neck region of adult nude mice. After 6–8 weeks the transplants and developed tumour regions were dissected, fixed in 4% paraformaldehyde and embedded in paraffin. For the evaluation of transplant differentiation, teratoma formation was defined as being when tissue representatives of at least all three germ layers were detected in the implants. Assessment of graft histology and function was performed by *in situ* hybridization, histochemistry with haemalaun and immunohistochemical methods.

Microsatellite profile analysis of DNA from all cells used in this study. Genomic DNA was extracted from the cell culture samples using the QIAamp MicroDNA Mini Kit (Qiagen). A selected set of DNA microsatellite markers was sent to the local Microarray Facility (Tübingen, Germany) for analysis. DNA samples were PCR amplified using a multiplex microsatellite marker panel (PowerPlex16, Promega). Amplicons were separated by capillary electrophoresis (ABI310, Applied Biosystems) and respective marker alleles were called using the PowerTyper macro in comparison with a co-analysed allelic ladder (Promega).

26. Schatten, G., Smith, J., Navara, C., Park, J. H. & Pedersen, R. Culture of human embryonic stem cells. *Nature Methods* **2**, 455–463 (2005).
27. Maltsev, V. A., Rohwedel, J., Hescheler, J. & Wobus, A. M. Embryonic stem cells differentiate *in vitro* into cardiomyocytes representing sinusnodal, atrial and ventricular cell types. *Mech. Dev.* **44**, 41–50 (1993).
28. Kehat, I. *et al.* Development of cardiomyocytes from human ES cells. *Methods Enzymol.* **365**, 461–473 (2003).
29. Ringe, J., Haupl, T. & Sittlinger, M. Mesenchymal stem cells for tissue engineering of bone and cartilage. *Med. Klin.* **98** (suppl. 2), 35–40 (2003).
30. Bielby, R. C., Boccaccini, A. R., Polak, J. M. & Buttery, L. D. *In vitro* differentiation and *in vivo* mineralization of osteogenic cells derived from human embryonic stem cells. *Tissue Eng.* **10**, 1518–1525 (2004).
31. Lumelsky, N. *et al.* Differentiation of embryonic stem cells to insulin-secreting structures similar to pancreatic islets. *Science* **292**, 1389–1394 (2001).
32. Segev, H., Fishman, B., Ziskind, A., Shulman, M. & Itskovitz-Eldor, J. Differentiation of human embryonic stem cells into insulin-producing clusters. *Stem Cells* **22**, 265–274 (2004).
33. Blyszczuk, P. *et al.* Embryonic stem cells differentiate into insulin-producing cells without selection of nestin-expressing cells. *Int. J. Dev. Biol.* **48**, 1095–1104 (2004).
34. Pollard, S. M., Conti, L., Sun, Y., Goffredo, D. & Smith, A. Adherent neural stem (NS) cells from fetal and adult forebrain. *Cereb. Cortex* **16** (suppl. 1), i112–i120 (2006).
35. Bibel, M. *et al.* Differentiation of mouse embryonic stem cells into a defined neuronal lineage. *Nature Neurosci.* **7**, 1003–1009 (2004).
36. Wittwer, C. T., Fillmore, G. C. & Hillyard, D. R. Automated polymerase chain reaction in capillary tubes with hot air. *Nucleic Acids Res.* **17**, 4353–4357 (1989).
37. Alexander, D. *et al.* Transcription factor Egr-1 activates collagen expression in immortalized fibroblasts or fibrosarcoma cells. *Biol. Chem.* **383**, 1845–1853 (2002).

CORRIGENDUM

doi:10.1038/nature08353

Generation of pluripotent stem cells from adult human testis

Sabine Conrad, Markus Renninger, Jörg Hennenlotter, Tina Wiesner, Lothar Just, Michael Bonin, Wilhelm Aicher, Hans-Jörg Bühring, Ulrich Mattheus, Andreas Mack, Hans-Joachim Wagner, Stephen Minger, Matthias Matzkies, Michael Reppel, Jürgen Hescheler, Karl-Dietrich Sievert, Arnulf Stenzl & Thomas Skutella

Nature 456, 344–349 (2008)

In this Letter, we omitted to disclose that the original patient consent forms to collect the material used to derive the pluripotent stem cells precluded distribution to third parties, as ruled on 8 December 2008, by the Ethics Commission of the School of Medicine and University Hospital Tübingen. Moreover, the Commission also ruled that the materials used to generate the lines were obtained from individuals who had signed consent forms that did not allow retention of the cells in culture for more than 3 years. We have now received broader consent of a few individuals permitting distribution of cells. These cells will be cultivated, and after quality tests they will be distributed to other scientists. The other individuals whose tissues were used in the study have opted for the terms of the initial consent, and thus their cell lines will be destroyed when the 3-year period runs out. We are now using the described protocols to generate new cell lines from individuals who have given explicit consent for distribution of lines to scientists.

ARTICLES

Structure of the intact PPAR- γ -RXR- α nuclear receptor complex on DNA

Vikas Chandra^{1*}, Pengxiang Huang^{1*}, Yoshitomo Hamuro², Srilatha Raghuram¹, Yongjun Wang³, Thomas P. Burris³ & Fraydoon Rastinejad¹

Nuclear receptors are multi-domain transcription factors that bind to DNA elements from which they regulate gene expression. The peroxisome proliferator-activated receptors (PPARs) form heterodimers with the retinoid X receptor (RXR), and PPAR- γ has been intensively studied as a drug target because of its link to insulin sensitization. Previous structural studies have focused on isolated DNA or ligand-binding segments, with no demonstration of how multiple domains cooperate to modulate receptor properties. Here we present structures of intact PPAR- γ and RXR- α as a heterodimer bound to DNA, ligands and coactivator peptides. PPAR- γ and RXR- α form a non-symmetric complex, allowing the ligand-binding domain (LBD) of PPAR- γ to contact multiple domains in both proteins. Three interfaces link PPAR- γ and RXR- α , including some that are DNA dependent. The PPAR- γ LBD cooperates with both DNA-binding domains (DBDs) to enhance response-element binding. The A/B segments are highly dynamic, lacking folded substructures despite their gene-activation properties.

The nuclear hormone receptors are a large family of transcription factors that directly bind and respond to ligands including steroids, thyroid hormone, retinoids, cholesterol by-products, lipids and haem^{1–4}. These receptors contain poorly conserved A/B regions that in some cases act as potent transcriptional activators, provide sites of protein phosphorylation or form direct interactions with other receptor domains or regulatory proteins⁵. A central and highly conserved DBD contains two zinc-binding sites and the architectural elements capable of sequence-specific binding to DNA^{5,6}. Hydrophobic molecules bind to the LBD, repositioning helix 12 into an active conformation that promotes the recruitment of co-regulators^{5,7}. The nuclear receptor coactivators, including members of the steroid receptor coactivator (SRC) family, contain LXXLL motifs that dock to LBDs^{8,9}. There have been multiple structural studies of nuclear receptors involving either the LBD or DBD fragments alone^{6,7,10–12}. However, there has been no successful visualization of any intact nuclear receptor. Consequently, there is little information about how different domains interface to impart complex physiological and pharmacological properties.

The PPARs, like many non-steroid members of the nuclear receptor family, function as obligate heterodimers with RXR¹³. PPAR- α , PPAR- β/δ and PPAR- γ are encoded separately, but have overlapping tissue expression patterns, and as a group coordinate the regulation of important metabolic pathways^{2,14,15}. These receptors control cellular processes including regulation of lipid and carbohydrate metabolism. PPAR- γ , the best-studied member of the family, is expressed in both white and brown adipocytes and regulates adipocyte differentiation, lipid storage and release¹⁶. One class of PPAR- γ ligands, the thiazolidinediones, which includes the drug rosiglitazone, are effective insulin sensitizers, and have been shown to improve glucose uptake and lower hyperglycaemia and hyperinsulinaemia^{17–20}. The PPARs are also potential therapeutic targets for atherosclerosis, inflammation and hypertension²¹.

To address how different domains of PPAR- γ and RXR- α polypeptides interact within and between their polypeptides, we performed a series of crystallographic studies using wild-type intact recombinant proteins. Both receptors' ligands, including rosiglitazone for PPAR- γ and 9-*cis*-retinoic acid for RXR- α , the DNA response element (PPRE) and coactivator peptides were included to examine how all the components of the complex interact. We further examined additional structures to determine if distinct ligands could modulate the multiple domain–domain interfaces. To this end we studied the crystal structures of two related complexes in which a partial agonist and a suicide inhibitor of PPAR- γ were used instead of rosiglitazone^{22,23}.

Overall organization

The structures we present here range in resolution from 3.1 to 3.2 Å, and were solved with single anomalous dispersion (SAD) phasing for the rosiglitazone structure, and molecular replacement for the GW9662 and BVT.13 complexes (Supplementary Table 1). Figure 1a, b shows the overall arrangement of the two receptors with respect to each other, their DNA-binding site, ligands and co-regulator (peptides) as seen in the rosiglitazone complex. As individual domains, the DBDs and LBDs of RXR- α , and the LBD of PPAR- γ , have folds highly similar to those previously reported for these isolated portions^{6,22,24}. The structure of the PPAR- γ DBD, although previously unknown, conforms to other DBDs as expected from the high level of sequence conservation^{6,25}. The rosiglitazone- and GW9662-containing structures belong to a different space group than the BVT.13-containing structure, and these crystal types display differences in packing. Nevertheless, the relative arrangements of the receptors, their domain–domain interactions, and the receptor interfaces with DNA and coactivators are similar in these crystal forms (Supplementary Fig. 1).

The hinge regions of nuclear receptors had not been previously seen beyond the short carboxy (C)-terminal extensions (CTEs) of

¹Department of Pharmacology, and Center for Molecular Design, University of Virginia Health System, 1300 Jefferson Park Avenue, Charlottesville, Virginia 22908-0735, USA. ²ExSAR Corporation, 11 Deer Park Drive, Suite 103, Monmouth Junction, New Jersey 08852, USA. ³Nuclear Receptor Biology Laboratory, Pennington Biomedical Research Center, Louisiana State University System, 6400 Perkins Road, Baton Rouge, Louisiana 70808, USA.

*These authors contributed equally to this work.

some DBDs²⁶. In PPAR- γ , the CTE forms a significant DNA interaction and is followed by two helical segments that reach the LBD (Fig. 1a, b). The RXR- α CTE forms one of the dimer contacts with PPAR- γ DBD, but is otherwise devoid of secondary structure and is more flexible. This flexibility may be an adaptive feature that allows RXR- α to be more promiscuous as a heterodimeric partner for numerous nuclear receptors, allowing their complexes to bind to direct repeats with multiple spacer sizes and half-site geometries. The PPAR- γ protein has its LBD and DBD closely positioned. In contrast, the RXR- α LBD and DBD are far apart from each other, with the space between them filled by the PPAR- γ LBD.

The PPAR- γ LBD has been the most exploitable drug site within the heterodimer, as ligands that bind to RXR- α are likely to inadvertently act on other RXR- α complexes. The PPAR- γ LBD is located prominently as the centrepiece of the complex, around which all other domains are positioned (Fig. 1). This arrangement leads to the consideration of whether some currently known or yet to be identified PPAR- γ ligands could impact other receptor domains. In contrast to the highly connected PPAR- γ LBD, the RXR- α LBD forms no contacts within the complex outside the PPAR- γ LBD. The two

receptor polypeptides are asymmetrically positioned with respect to each other, whether viewed as intact polypeptides or as individual domains (Figs 1 and 2b).

Binding to PPRE DNA

The PPRE element we used contains two copies of the consensus half-site AGGTCA, separated by a single base pair, forming the so-called DR1 element (Fig. 2A)²⁷. This element contains an additional AAACCT motif positioned 5' to the DR1²⁷. The two half-sites are distinguishable by their 5' and 3' positioning on the DR1 (Fig. 2a, b). PPAR- γ and RXR- α display half-site selective binding that results in a polar arrangement of their DBDs, with PPAR- γ residing upstream of RXR- α . This polarity is in agreement with a previous biochemical study, and is seen in all three crystal structures described here²⁷.

A major determinant of the polarity is the PPAR- γ CTE, which interacts significantly with the 5' flanking sequence (described below). A second and more subtle determinant of the polarity is the requirement that both receptors form productive interactions between their DBDs. PPAR- γ and RXR- α DBDs interact within the minor groove of the spacer (Fig. 2b). The DNA dependence of DBD–DBD contacts is clear as the two proteins bury a modest 30 Å² surface area between them. This size is negligible compared with standard dimer interfaces formed between protein pairs. The spacer DNA minor groove shelters the highly polar side chains of the interacting residues (Asn 160 from PPAR- γ , and Arg 209 and Gln 206 of RXR- α) from solvent. One interpretation of this arrangement is that the PPRE DNA allosterically contributes to its own binding by directly bridging the two receptor DBDs. We previously described the crystallographic arrangement of RXR–RAR on DR1, based on their isolated DBDs²⁸. The observed polarity and DBD–DBD contacts are precisely the same for PPAR–RXR, as are the residues involved in these dimer interfaces (Supplementary Fig. 2).

Both DBDs have α -helices directly in register with the AGGTCA sequences (Supplementary Fig. 3). Hydrogen-bonding contacts are made between the DBD amino acids and the major groove features of the half-sites. PPAR- γ makes more base and phosphate backbone interactions than RXR- α (Supplementary Figs 3 and 4). The PPAR- γ

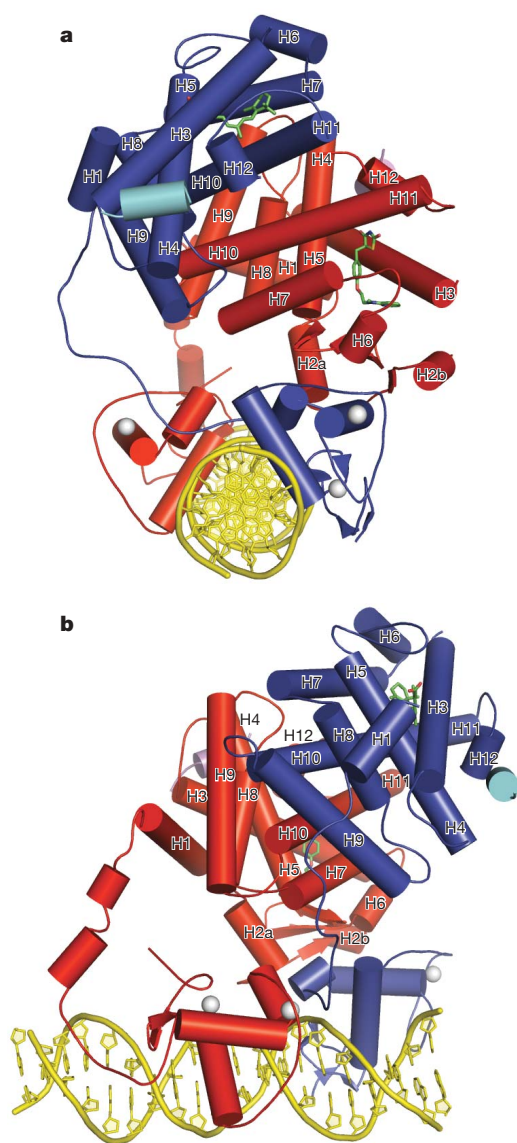


Figure 1 | Overall structure of the PPAR- γ -RXR- α complex on PPRE. **a, b,** Orthogonal views are shown in which RXR- α is blue and PPAR- γ is red. The ligands rosiglitazone and 9-*cis*-retinoic acid are shown in green, the Zn(II) ions are white, and the coactivator LXXLL peptides are in light blue and purple.

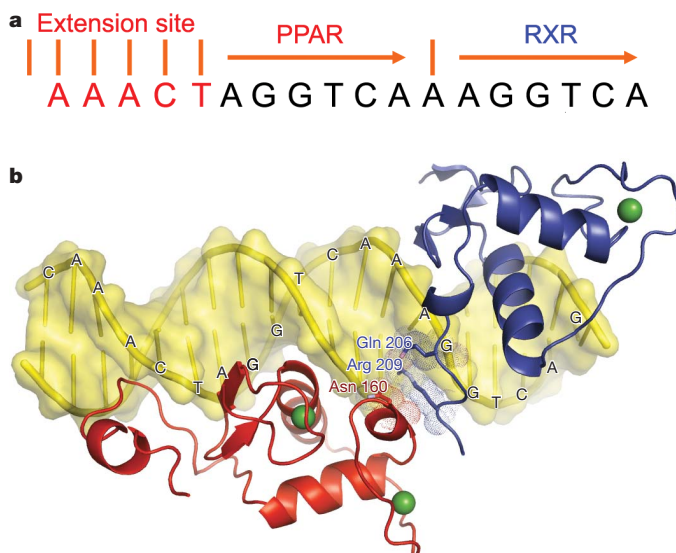
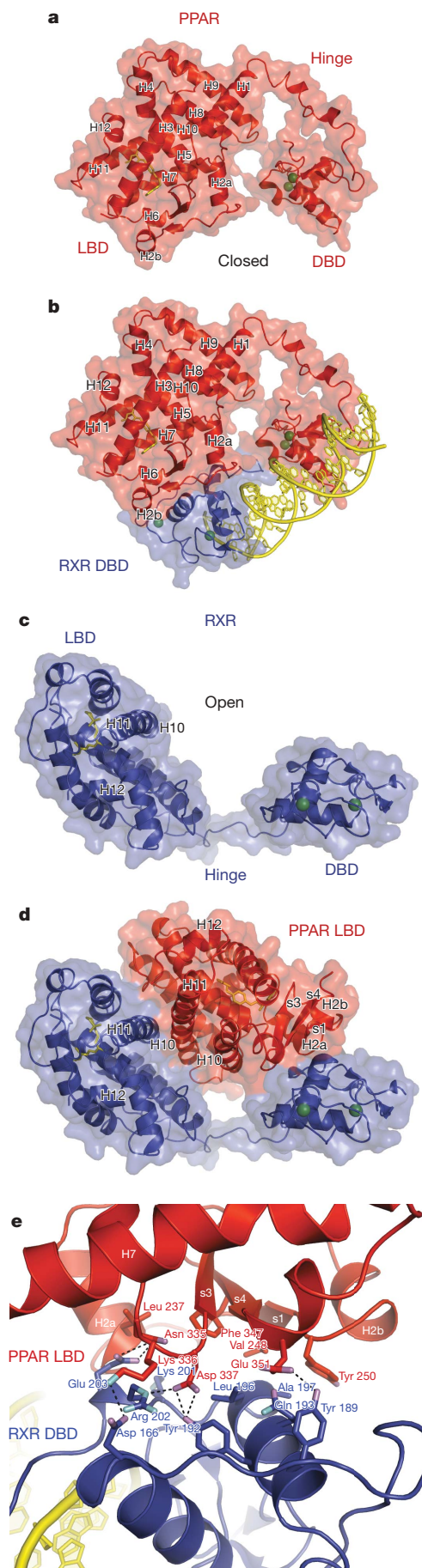


Figure 2 | DNA-binding features of the complex. **a,** The sequence of PPRE element containing direct repeats of the canonical AGGTCA half-site separated by one base pair (DR1) together with the upstream sequence AAACCT specificity element. **b,** The DBDs and their C-terminal extensions are shown for PPAR- γ (red) and RXR- α (blue). These domains form a polar, DNA-dependent head-to-tail interaction through side chains that converge over the spacer element.



hinge region (CTE portion) makes an extensive DNA interaction as well, binding to the upstream AAAC element (Supplementary Fig. 4). This interaction is similar to that of the Rev-Erb CTE with its AAAC-binding sequence upstream of its DR2-binding site^{26,29}. Both PPAR and Rev-Erb deeply embed the tetrapeptide RFGR from their CTEs into this DNA minor groove (Supplementary Fig. 4). DNA sequences are not clearly distinguishable from their minor groove side, yet in the case of PPREs the AAAC sequence is well conserved²⁷. We showed previously that this conserved sequence provides the DNA deformability and minor groove distortions that allow recognition by RFGR sequences²⁹.

Cooperation between receptor domains

The DBD–DBD interface is only one of the three distinct heterodimerization surfaces between the receptors. Additionally, a dimer interface is formed by the two LBDs (Fig. 1a, b)²⁴. This interface is DNA independent and buries 2,160 Å² of solvent-accessible surface, using helices 7, 9 and 10 of each receptor. A third, previously unknown, heterodimerization interface is also observed between the PPAR-γ LBD and the DBD CTE region of RXR-α. The DNA dependence of this interaction is suggested in Fig. 3b, e. Electron density of all three of the heterodimerization interfaces is shown in Supplementary Figures 5–7.

Figures 3c, d demonstrate how both the DBD and LBD of RXR-α are intersected by the PPAR-γ LBD. A feature unique to the LBD of PPARs is their β-strand elements S1, S2, S3 and S4, which face the RXR-α DBD (Figs 3b, d, e). The linchpin of this interaction appears to be PPAR-γ LBD's Phe 347, which resides in strand S4 and forms hydrophobic interactions with RXR-α DBD (Fig. 3e). As the amino acids shown in Fig. 3e are conserved among both PPARs and RXRs, other PPAR family members are expected to form the same interface.

The PPAR-γ LBD contacts both receptor DBDs to stabilize their DR1 binding (Fig. 3b). The DNA-binding properties of intact nuclear receptors are typically of higher affinity than those of their isolated DBDs³⁰. The current structure shows how regions outside the DBD can contribute to DNA recognition and affinity by interacting with DBDs. The PPAR-γ LBD also brings its own positively charged surface near the negatively charged DNA phosphodiester backbone. An example is Lys 240 from the PPAR-γ LBD, which has its basic side chain positioned in close proximity to the DNA backbone, indirectly contributing to favourable electrostatic interactions with DNA.

To test the idea that the PPAR-γ LBD interface with RXR-α DBD is contributing to the DNA binding of the complex, we made a point mutation at a key hydrophobic residue in PPAR-γ LBD, Phe 347. This residue is near the surface of the LBD, and distant from both the ligand-binding pocket and coactivator docking sites (Fig. 3e). Our DNA binding assays show that the Phe347Ala-mutant PPAR-γ negatively affects PPRE binding (Supplementary Figs 8 and 9). The weakened ability to bind DNA is further manifested in the receptors' inability to activate transcription in response to rosiglitazone (Supplementary Fig. 10). These data demonstrate that a residue positioned on the surface of a receptor LBD can strongly influence DNA binding, confirming the ability of the PPAR-γ LBD, through domain–domain interactions, to affect the receptors' DNA-binding properties.

A comparison of the ligand-binding pockets of the intact receptors with previous structurally defined LBDs alone did not reveal any unexpected differences. Functional PPREs almost exclusively consist

Figure 3 | Domain–domain interactions involving the PPAR-γ LBD. **a**, The PPAR-γ LBD is closely positioned to the PPAR-γ DBD. **b**, The PPAR-γ LBD (red) interfaces with both receptor DBDs to enhance their DNA association. RXR DBD is in blue. **c**, The LBD and DBD of RXR-α, do not interact. **d**, PPAR-γ LBD bisects both RXR-α domains. The proximity of the PPAR-γ ligand to both RXR-α domains can be seen. **e**, The side-chain interactions between PPAR-γ LBD and RXR-α DBD in the rosiglitazone complex. The PPAR-γ LBD residue Phe 347 lies at the centre of the interface.

of DR1 sites³¹. Other RXR heterodimers, which can use multiple response elements, are likely to show different types of domain–domain interaction, and possibly different modes of ligand binding, when bound to their different target DNAs. For example, the RAR–RXR complex displays a distinctly different behaviour for ligand binding and gene activation on DR5 versus DR1 elements³².

Effect of PPAR- γ ligands on receptor organization

We considered whether different ligands could affect the domain–domain interactions. We selected two other PPAR- γ ligands that we believed would have distinct properties from rosiglitazone. The first was the antagonist GW9662, which can act as a suicide inhibitor of PPAR- γ ²³. A second ligand we used was BVT.13, which in its bound position was expected to have a significant reach into the unique β -strand region of the PPAR- γ LBD, the region strongly coupled to the RXR- α DBD²². In all three crystal structures, we confirmed by

electron density that these ligands were occupying the pocket (Fig. 4a).

As a group, rosiglitazone, GW9662 and BVT.13 gave rise to a 'Y'-shaped pocket (Fig. 4b). GW9662, which has been shown biochemically to be capable of forming a covalent bond with a cysteine in the PPAR- γ LBD pocket, resides in the pocket without having made the covalent bond in the intact PPAR- γ (Fig. 4c). Nevertheless, the positioning of the nitro-aryl chloride portion of GW9662 shows this ligand is poised to form a covalent adduct, given the proximity of its displaceable chloride group to the cysteine sulphur. The covalent bond could in principle form in a small population of PPAR- γ receptors that were not captured in our crystallization condition. BVT.13 has a longer reach and is inserted more directionally towards the β -strand segments of the PPAR- γ LBD (Fig. 4a, b). A previous study showed that this ligand could also dramatically influence the dynamics of the β -strand elements of PPAR- γ LBD²². With both

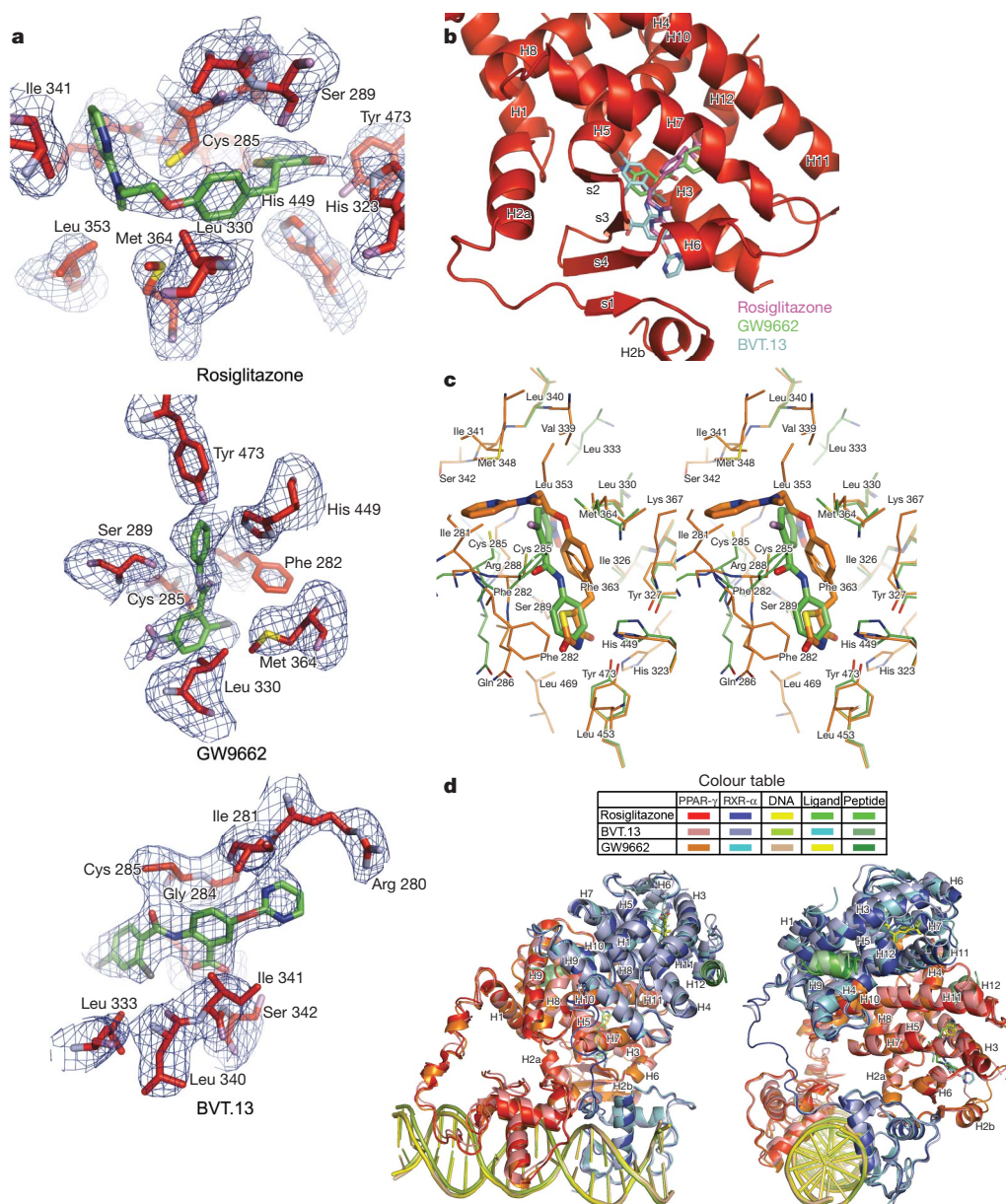


Figure 4 | Effect of different ligands on the PPAR- γ -RXR- α complex. **a**, Electron density ($F_o - F_c$) omit maps around the PPAR- γ ligands rosiglitazone, BVT.13 and GW9662 from the their respective complexes. **b**, The three PPAR- γ ligands as a group occupy distinct portions of the receptor's Y-shaped pocket. **c**, Comparison of the binding mode of

rosiglitazone with GW9662 within the PPAR- γ pocket. GW9662 does not form a covalent bond with Cys 285. **d**, Superposition of crystal structures obtained with the different PPAR- γ ligands showing minor variations in overall domain organization. The colours are indicated for each complex.

the GW9662 and BVT.13 ligands, the overall conformation of the two receptors, including all of the domain–domain, receptor–DNA, and receptor–coactivator interactions, were not found significantly altered from those seen in the rosiglitazone complex (Fig. 4d).

Dynamic properties of PPAR- γ

The A/B segment of PPAR- γ is a potent transcriptional activator, contains a phosphorylation site and can interact with other proteins^{33–36}. However, the A/B segments could not be visualized in any of the three complexes, presumably because of their high mobility. We examined their dynamic properties using amide hydrogen/deuterium exchange mass spectrometry (H/D-Ex)^{37,38}. The H/D exchange rate is significantly slower in structured regions than unstructured portions. The H/D exchange levels of the intact

PPAR- γ with no ligands, with rosiglitazone or with GW9662 were measured. The H/D exchange levels of a third complex were also measured which simultaneously contained rosiglitazone, RXR- α , 9-*cis*-retinoic acid and PPRE DR1, similar to our crystallographic samples but lacking peptide elements. The H/D-Ex analysis allowed us to judge the relative dynamic properties of segments along the PPAR- γ protein.

Figure 5a shows that DNA protected the amino-acid region 170–193 from exchange, the portion of PPAR- γ that inserts into the minor groove of the 5' DR1 extension. The amino-acid region 434–442 in helix 10 and region 333–340 also showed slower exchange compared with the PPAR- γ protein without RXR- α . The amino-acid region 434–442 is involved in the LBD–LBD heterodimerization and region 333–340 interacts with RXR- α DBD (see residues 335–337 in Fig. 3e).

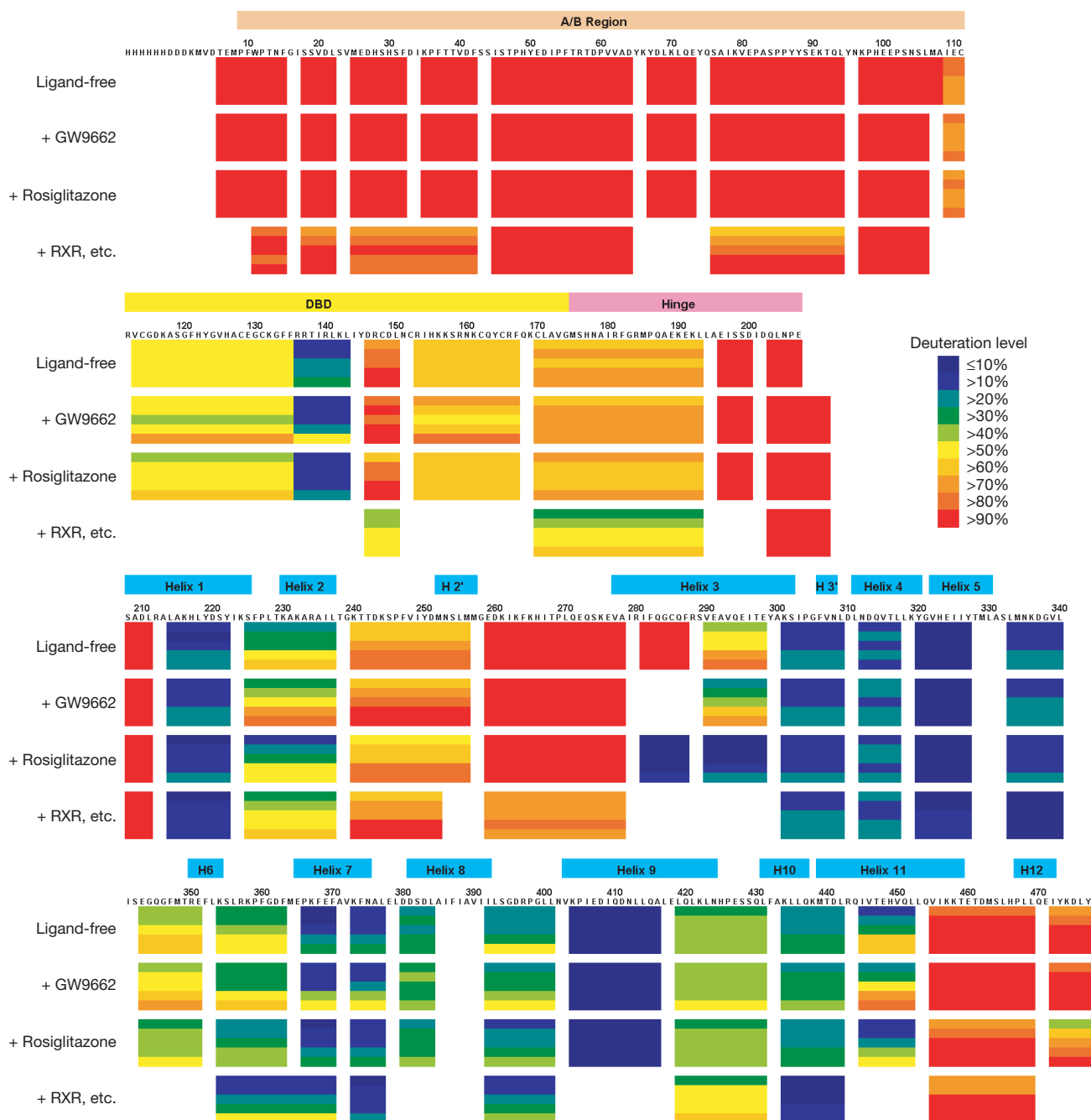


Figure 5 | The dynamic features of the receptors. H/D-Ex data for intact PPAR- γ . From top, in a ligand-free state, with rosiglitazone, with GW9662, and with all of rosiglitazone, RXR- α , 9-*cis*-retinoic acid and PPRE DNA (RXR, etc.). Each horizontal colour block represents an analysed peptic

fragment. The deuteration level of each peptide at each time point is colour-coded as in the insert. Each block contains five time points (from top, 15, 50, 150, 500, and 1,500 s).

Importantly, in all samples, the most rapidly exchanging region was the A/B portion. Sequence alignments further demonstrate that these receptors lack significant stretches of hydrophobic residues and/or a meaningful level of amino-acid conservation in their A/B regions to suggest the existence of a folded motif (Supplementary Figs 11 and 12).

Coactivator interactions

Both RXR- α and PPAR- γ assume the so-called active conformation on their LBD in all three structures. This conformation consists of both receptors having their helix 12 appropriately positioned by ligands to facilitate the docking of coactivator LXXLL motifs (Fig. 1a, b). In the context of the entire complex, the location of the coactivator peptides is distantly isolated from all of the other protein–protein interaction sites (Fig. 1). This positioning suggests no direct modulation of coactivator binding from other receptor domains. The PPARE is also very distant from the coactivator-binding site. Therefore, to recruit promoter-specific co-regulators, PPAR binding sites may be required to cooperate with other specific promoter elements that are positioned nearby.

A so-called phantom ligand effect has also been described in some RXR complexes, where an RXR ligand can cause dissociation of co-repressors and recruitment of coactivators at the non-covalently bound subunit of the heterodimer³⁹. The current work does not provide a direct explanation for the phantom ligand effect or for receptor permissivity; instead, it shows two individual LXXLL motifs interacting independently at the PPAR- γ and RXR- α LBDs. In the case of RXR–RAR heterodimers, it has been shown that synergy between their ligands, resulting in enhanced transcription, can be created by the increased interaction efficiency of a single p160 protein⁴⁰.

Discussion

The importance of the PPARs as a nuclear receptor subfamily is underscored by their validated roles in human disease and the proven beneficial role of some PPAR ligands for patients with type 2 diabetes. As many other non-steroid receptors also form heterodimers with RXR, the current structural analysis suggests the possibility of many different types of domain–domain interaction both within a single polypeptide and through the heterodimeric partners. The PPAR- γ –RXR- α polypeptides reside in a polar arrangement set by the 5' extension of the DR1 and the lone spacer base pair. PPAR- γ occupies most of the DR1 including its specific 5' element (Supplementary Fig. 13). The DNA element largely dictates the type of domain–domain interactions in the complex, by predisposing both receptors to a unique polarity and spatial arrangement.

The observed architecture of the DBD and LBD receptor domains, in the context of the current structures, does not differ significantly from the same domains crystallized in isolation^{12,24,41}. However, the domain–domain interactions involving intact nuclear receptors are defined fully here for the first time. The PPAR- γ LBD lies at the centre of the organized complex and is juxtaposed against every ordered domain from both proteins. Three distinct heterodimerization interfaces are formed between PPAR- γ and RXR- α . The interconnected nature of both receptors is intimately linked to their mode of binding to the PPARE DNA. Moreover, the PPAR- γ LBD assists both DBDs in their binding to DNA. This observation is supported by the finding that a single point mutation in the PPAR- γ LBD can severely impact DNA binding by the heterodimer. The structures also allow for a better understanding of PPAR- γ missense mutations associated clinically with partial lipodystrophy, insulin resistance and colon cancer (Supplementary Table 2).

We found that the A/B domains are intrinsically flexible as they were not observable in the crystal structure. These receptor portions are also not conserved in sequence. The intrinsic flexibility in this receptor portion of PPAR- γ was independently verified by H/D-Ex. Examination of several different crystal structures of PPAR- γ –RXR- α complexes indicated that three distinct PPAR- γ ligands did not substantially impact the ordering of the A/B regions or alter the domain–

domain interactions within the complexes. A summary of all the critical sites of interactions on PPAR- γ , RXR- α and the PPARE is presented in Supplementary Figures 11–13. We expect that the other PPARs could form a similar arrangement with RXRs on DR1 PPAREs, as the major sites of domain–domain interactions are well conserved in amino-acid sequence.

Owing to their direct involvement in many disease pathways, most of the nuclear receptors have been the subject of compound screening assays to find selective small-molecule modulators that bind to the LBD pockets. For practical reasons, such efforts at drug discovery have mainly relied on isolated LBDs, not intact receptor homo- or heterodimers on DNA. Moreover, the major readout for compound screening assays has been co-regulator recruitment properties. These studies have been guided by the assumption that different nuclear receptor ligands can impart their gene-selective actions strictly through their differential impact on co-regulator affinities. Given that co-regulator availability differs in cell types, the altered affinities are thought to give rise to selective modulator properties in these ligands.

The current structural examination shows that the LBD of a receptor can be tightly coupled to other receptor segments both within its own polypeptide and with the heterodimeric partner. The LBD can also have an important function in modulating DNA binding depending on the manner in which the complex is organized. It remains plausible that some nuclear receptor ligands could provide their selective gene modulation properties by imparting a graded set of response element affinities, and not through coactivator affinities alone. More sophisticated small-molecule screening and crystallization efforts are needed that use fully assembled and intact nuclear receptors to test this concept properly.

METHODS SUMMARY

Expression, purification and crystallization. The hRXR- α (NP_002948) and hPPAR- γ (NP_619726) proteins were expressed using pET46 Ek/LIC vectors in BL21(DE3) *Escherichia coli* cells (Novagen). Cells were induced with 1.0 mM IPTG at 17 °C overnight, collected and lysed in 20 mM Tris, pH 7.5, 500 mM NaCl, 30 mM imidazole and 10% glycerol. Proteins were purified on His-Bind resin (Novagen) and on SP-Sephacrose columns (GE Healthcare). Ligands, coactivator-peptide (EKHKILHRLQDSY) and DNA were added. The oligonucleotide strands 5'-GCAAACTAGGTCAAAGGTCAG-3' and 5'-CTGACCT-TTGACCTAGTTTGC-3' were prepared as described previously²⁶. Crystals were grown with 15–18% PEG 3350, 25 mM MgCl₂, 100 mM NH₄Cl, 5 mM dithiothreitol, 0.1 M MES, pH 6.5, at 4 °C.

Data collection and structure determination. Diffraction data were collected at 100 K at the Argonne National Laboratory beamline SER-CAT 20ID and processed with HKL-3000 (ref. 42). The rosiglitazone-containing structure was solved using phases obtained from a SAD experiment. Both Refmac5 (CCP4) and CNS 1.2 were used for initial structure refinement, during which restraints were placed on DNA geometry^{43–45}. The rosiglitazone structure was used to solve the BVT.13 and GW9662 structures by molecular replacement using PHASER⁴⁶.

Hydrogen/deuterium exchange experiments. An exchange reaction was initiated by mixing 10 μ l of protein stock solution (with or without ligand and/or DNA) with 10 μ l of D₂O. After incubation at 4 \pm 1 °C for a predetermined time (15–1,500 s), the H/D-exchange reaction was quenched by shifting to pH 2.5 with 30 μ l of 8 M urea and 1 M TCEP at 1 °C. The quenched solution was immediately pumped over a pepsin column (104 μ l bed volume)³⁸. Peptide fragments were separated by a C18 column. Mass spectrometric analyses were performed with a Thermo Fisher LCQ instrument, and the centroids of probe peptide isotopic envelopes were determined using a program developed in-house⁴⁷. The corrections for back-exchange were made using methods described previously⁴⁸.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 1 May; accepted 10 September 2008.

Published online 29 October 2008; corrected 20 November 2008 (details online).

1. Nagy, L. & Schwabe, J. W. Mechanism of the nuclear receptor molecular switch. *Trends Biochem. Sci.* **29**, 317–324 (2004).
2. Sonoda, J., Pei, L. & Evans, R. M. Nuclear receptors: decoding metabolic disease. *FEBS Lett.* **582**, 2–9 (2008).

3. Raghuram, S. *et al.* Identification of heme as the ligand for the orphan nuclear receptors REV-ERB α and REV-ERB β . *Nature Struct. Mol. Biol.* **14**, 1207–1213 (2007).
4. Yin, L. *et al.* Rev-erb α , a heme sensor that coordinates metabolic and circadian pathways. *Science* **318**, 1786–1789 (2007).
5. Bain, D. L., Heneghan, A. F., Connaghan-Jones, K. D. & Miura, M. T. Nuclear receptor structure: implications for function. *Annu. Rev. Physiol.* **69**, 201–220 (2007).
6. Khorasanizadeh, S. & Rastinejad, F. Nuclear-receptor interactions on DNA-response elements. *Trends Biochem. Sci.* **26**, 384–390 (2001).
7. Wurtz, J. M. *et al.* A canonical structure for the ligand-binding domain of nuclear receptors. *Nature Struct. Biol.* **3**, 87–94 (1996).
8. McKenna, N. J. & O'Malley, B. W. Combinatorial control of gene expression by nuclear receptors and coregulators. *Cell* **108**, 465–474 (2002).
9. Heery, D. M., Kalkhoven, E., Hoare, S. & Parker, M. G. A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature* **387**, 733–736 (1997).
10. Li, Y., Lambert, M. H. & Xu, H. E. Activation of nuclear receptors: a perspective from structural genomics. *Structure* **11**, 741–746 (2003).
11. Greschik, H. & Moras, D. Structure-activity relationship of nuclear receptor-ligand interactions. *Curr. Top. Med. Chem.* **3**, 1573–1599 (2003).
12. Rastinejad, F. Retinoid X receptor and its partners in the nuclear receptor family. *Curr. Opin. Struct. Biol.* **11**, 33–38 (2001).
13. Chawla, A., Repa, J. J., Evans, R. M. & Mangelsdorf, D. J. Nuclear receptors and lipid physiology: opening the X-files. *Science* **294**, 1866–1870 (2001).
14. Chen, F., Law, S. W. & O'Malley, B. W. Identification of two mPPAR related receptors and evidence for the existence of five subfamily members. *Biochem. Biophys. Res. Commun.* **196**, 671–677 (1993).
15. Lazar, M. A. PPAR γ , 10 years later. *Biochimie* **87**, 9–13 (2005).
16. Lehrke, M. & Lazar, M. A. The many faces of PPAR γ . *Cell* **123**, 993–999 (2005).
17. Lehmann, J. M. *et al.* An antidiabetic thiazolidinedione is a high affinity ligand for peroxisome proliferator-activated receptor γ (PPAR γ). *J. Biol. Chem.* **270**, 12953–12956 (1995).
18. Olefsky, J. M. & Saltiel, A. R. PPAR γ and the treatment of insulin resistance. *Trends Endocrinol. Metab.* **11**, 362–368 (2000).
19. Staels, B. PPAR agonists and the metabolic syndrome. *Therapie* **62**, 319–326 (2007).
20. Willson, T. M. *et al.* The structure-activity relationship between peroxisome proliferator-activated receptor γ agonism and the antihyperglycemic activity of thiazolidinediones. *J. Med. Chem.* **39**, 665–668 (1996).
21. Berger, J. & Wagner, J. A. Physiological and therapeutic roles of peroxisome proliferator-activated receptors. *Diabetes Technol. Ther.* **4**, 163–174 (2002).
22. Bruning, J. B. *et al.* Partial agonists activate PPAR γ using a helix 12 independent mechanism. *Structure* **15**, 1258–1271 (2007).
23. Leesnitzer, L. M. *et al.* Functional consequences of cysteine modification in the ligand binding sites of peroxisome proliferator activated receptors by GW9662. *Biochemistry* **41**, 6640–6650 (2002).
24. Gampe, R. T. Jr *et al.* Asymmetry in the PPAR γ /RXR α crystal structure reveals the molecular basis of heterodimerization among nuclear receptors. *Mol. Cell* **5**, 545–555 (2000).
25. Rastinejad, F., Perlmann, T., Evans, R. M. & Sigler, P. B. Structural determinants of nuclear receptor assembly on DNA direct repeats. *Nature* **375**, 203–211 (1995).
26. Zhao, Q., Khorasanizadeh, S., Miyoshi, Y., Lazar, M. A. & Rastinejad, F. Structural elements of an orphan nuclear receptor-DNA complex. *Mol. Cell* **1**, 849–861 (1998).
27. Ijpenberg, A., Jeannin, E., Wahli, W. & Desvergne, B. Polarity and specific sequence requirements of peroxisome proliferator-activated receptor (PPAR)/retinoid X receptor heterodimer binding to DNA. A functional analysis of the malic enzyme gene PPAR response element. *J. Biol. Chem.* **272**, 20108–20117 (1997).
28. Rastinejad, F., Wagner, T., Zhao, Q. & Khorasanizadeh, S. Structure of the RXR-RAR DNA-binding complex on the retinoic acid response element DR1. *EMBO J.* **19**, 1045–1054 (2000).
29. Sierk, M. L., Zhao, Q. & Rastinejad, F. DNA deformability as a recognition feature in the reverber response element. *Biochemistry* **40**, 12833–12843 (2001).
30. Ozers, M. S. *et al.* Equilibrium binding of estrogen receptor with DNA using fluorescence anisotropy. *J. Biol. Chem.* **272**, 30405–30411 (1997).
31. Desvergne, B. & Wahli, W. Peroxisome proliferator-activated receptors: nuclear control of metabolism. *Endocr. Rev.* **20**, 649–688 (1999).
32. Kurokawa, R. *et al.* Regulation of retinoid signalling by receptor polarity and allosteric control of ligand binding. *Nature* **371**, 528–531 (1994).
33. Werman, A. *et al.* Ligand-independent activation domain in the N terminus of peroxisome proliferator-activated receptor γ (PPAR γ). Differential activity of PPAR γ -1 and -2 isoforms and influence of insulin. *J. Biol. Chem.* **272**, 20230–20235 (1997).
34. Adams, M., Reginato, M. J., Shao, D., Lazar, M. A. & Chatterjee, V. K. Transcriptional activation by peroxisome proliferator-activated receptor gamma is inhibited by phosphorylation at a consensus mitogen-activated protein kinase site. *J. Biol. Chem.* **272**, 5128–5132 (1997).
35. Castillo, G. *et al.* An adipogenic cofactor bound by the differentiation domain of PPAR γ . *EMBO J.* **18**, 3676–3687 (1999).
36. Shao, D. *et al.* Interdomain communication regulating ligand binding by PPAR- γ . *Nature* **396**, 377–380 (1998).
37. Maier, C. S. & Deinzer, M. L. Protein conformations, interactions, and H/D exchange. *Methods Enzymol.* **402**, 312–360 (2005).
38. Hamuro, Y. *et al.* Rapid analysis of protein structure and dynamics by hydrogen/deuterium exchange mass spectrometry. *J. Biomol. Tech.* **14**, 171–182 (2003).
39. Schulman, I. G., Li, C., Schwabe, J. W. & Evans, R. M. The phantom ligand effect: allosteric control of transcription by the retinoid X receptor. *Genes Dev.* **11**, 299–308 (1997).
40. Germain, P., Iyer, J., Zechel, C. & Gronemeyer, H. Co-regulator recruitment and the mechanism of retinoic acid receptor synergy. *Nature* **415**, 187–192 (2002).
41. Nolte, R. T. *et al.* Ligand binding and co-activator assembly of the peroxisome proliferator-activated receptor- γ . *Nature* **395**, 137–143 (1998).
42. Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. HKL-3000: the integration of data reduction and structure solution—from diffraction images to an initial model in minutes. *Acta Crystallogr. D* **62**, 859–866 (2006).
43. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D* **53**, 240–255 (1997).
44. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
45. Brunger, A. T. Version 1.2 of the Crystallography and NMR system. *Nature Protoc.* **2**, 2728–2733 (2007).
46. McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D* **63**, 32–41 (2007).
47. Hamuro, Y. *et al.* Hydrogen/deuterium-exchange (H/D-Ex) of PPAR γ LBD in the presence of various modulators. *Protein Sci.* **15**, 1–10 (2006).
48. Zhang, Z. & Smith, D. L. Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation. *Protein Sci.* **2**, 522–531 (1993).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank K. S. Molnar, S. J. Tuske and S. J. Coales for providing assistance with the H/D-Ex studies; M. Chruszcz and W. Minor for assistance with diffraction data processing and analysis; P. Rogers for analysis of mutant receptor expression; and P. Griffin for providing BVT.13.

Author Contributions V.C. expressed, purified and crystallized the samples, and with P.H. collected data and solved the structure. Y.H. performed the H/D-Ex work. S.R. provided the expression systems. Y.W. and T.P.B. performed the electrophoretic mobility shift assay and transcription reporter assays. F.R. supervised the work and wrote the manuscript.

Author Information Atomic coordinates and structure factors for the reported crystal structures are deposited in the Protein Data Bank under accession numbers 3DZY, 3DZU and 3E00. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to F.R. (fr9c@virginia.edu).

METHODS

Expression, purification and crystallization. The hRXR- α (NP_002948) and hPPAR- γ (NP_619726) were expressed as His-tagged proteins from pET46 Ek/LIC vector in BL21(DE3) *E. coli* cells (Novagen). The seleno-methionyl derivative of hRXR- α was prepared by expression in B834(DE3) *E. coli* grown in minimal medium supplemented with Seleno-Met (Sigma). Cells were induced with 1.0 mM IPTG at 17 °C overnight, harvested and lysed in 20 mM Tris, pH 7.5, 500 mM NaCl, 30 mM imidazole and 10% glycerol. Each protein was purified on His-Bind resins (Novagen) and subsequently on SP-Sepharose columns (GE Healthcare), and their concentrations estimated based on calculated extinction values. 9-*cis*-retinoic acid (Sigma) and rosiglitazone (Cayman Chemicals) were added at 7 \times molar excess to RXR and PPAR, respectively. Synthetic coactivator peptide (EKHKILHRLQDSY) was added in 2 \times molar ratio to each receptor. RXR and PPAR were then combined in 1:1 molar ratio and subsequently mixed with DNA oligonucleotide (DR1) at 1.5 \times molar ratio. The oligonucleotide strands containing the essential segments 5'-GCAAACTAGGTCAAAGTCAG-3' and 5'-CTGACCTTTGACCTAGTTTGC-3' were purified and annealed as described previously²⁶. The heterodimeric complex was purified on Superdex 200 (GE Healthcare) gel filtration column. The complex was concentrated to 3.0 mg ml⁻¹ and crystals were grown using the hanging drop method against a reservoir of 15–18% PEG 3350, 25 mM MgCl₂, 100 mM NH₄Cl, 5 mM dithiothreitol, 0.1 M MES, pH 6.5, at 4 °C. Crystals were streaked through a cryosolvent solution containing the reservoir solution supplemented with 34% glycerol and flash frozen in liquid nitrogen.

Data collection and structure determination. Diffraction data were collected on frozen crystals at 100 K using the Argonne National Laboratory at beamline SER-CAT 20ID. These data were integrated and scaled using HKL-3000 (ref. 42). The rosiglitazone-containing structure was initially solved, wherein phases were obtained using a SAD-phased electron-density map derived from the 3.1-Å selenomethionine-RXR-containing complex. Selenium sites were identified and refined using SHELXD, MLPHARE and DM features of HKL-3000. Both Refmac5 (CCP4) and CNS 1.2 were used for initial structure refinement, during which restraints were placed on DNA geometry^{43–45}. Restraints were released in the later stages of refinement. The 3.1-Å refined rosiglitazone complex structure was used to solve the BVT.13 and GW9662 structures by molecular replacement using the program PHASER⁴⁶. All data collection and refinement statistics are listed in Supplementary Table 1.

Hydrogen/deuterium exchange experiments. An exchange reaction was initiated by mixing 10 μ l of protein stock solution (18 μ M PPAR- γ \pm 144 μ M ligand \pm DNA in 50 mM Tris, 200 mM NaCl, pH 8.0) with 10 μ l of D₂O. After incubation at 4 \pm 1 °C for a predetermined time (15, 50, 150, 500 or 1,500 s), the H/D-exchange reaction was quenched by shifting to pH 2.5 with 30 μ l of 8 M urea and 1 M TCEP at 1 °C. The quenched solution was immediately pumped over a pepsin column (104 μ l bed volume) filled with porcine pepsin (Sigma) immobilized on Poros 20 AL media at 30 mg ml⁻¹ according to the manufacturer's instructions, with 0.05% TFA (200 μ l min⁻¹) for 3 min with contemporaneous collection of proteolytic products by a trap column (4 ml bed volume). Subsequently, the peptide fragments were eluted from the trap column and separated by C18 column (Magic C18, Michrom BioResources) with a linear gradient of 15% solvent B to 40% solvent B over 23 min (solvent A, 0.05% TFA in water; solvent B, 80% acetonitrile, 20% water, 0.01% TFA; flow rate 5–10 μ l min⁻¹). Mass spectrometric analyses were performed with an LCQ mass spectrometer (Thermo Fisher) with capillary temperature at 200 °C.

The centroids of probe peptide isotopic envelopes were measured using a program developed in-house in collaboration with Sierra Analytics⁴⁷. The corrections for back-exchange were made using methods described previously⁴⁸.

$$\text{Deuteration level (\%)} = [(m(P) - m(N))/(m(F) - m(N))] \times 100$$

$$\text{Deuterium incorporation (\#)} = [(m(P) - m(N))/(m(F) - m(N))] \times \text{MaxD}$$

where $m(P)$, $m(N)$ and $m(F)$ are the centroid value of partly deuterated peptide, non-deuterated peptide and fully deuterated peptide, respectively. MaxD is the maximum deuterium incorporation calculated by subtracting the number of prolines in the third or later amino acid and two from the number of amino acids in the peptide of interest (assuming the first two amino acids cannot retain deuterons⁴⁹). For PPAR- γ full length, the average deuterium recovery of a fully deuterated sample $((m(F) - m(N))/\text{MaxD})$ was 66%.

49. Bai, Y., Milne, J. S., Mayne, L. & Englander, S. W. Primary structure effects on peptide group hydrogen exchange. *Proteins* 17, 75–86 (1993).

Identification of Holliday junction resolvases from humans and yeast

Stephen C. Y. Ip^{1*}, Ulrich Rass^{1*}, Miguel G. Blanco^{1*}, Helen R. Flynn², J. Mark Skehel² & Stephen C. West¹

Four-way DNA intermediates, also known as Holliday junctions, are formed during homologous recombination and DNA repair, and their resolution is necessary for proper chromosome segregation. Here we identify nucleases from *Saccharomyces cerevisiae* and human cells that promote Holliday junction resolution, in a manner analogous to that shown by the *Escherichia coli* Holliday junction resolvase RuvC. The human Holliday junction resolvase, GEN1, and its yeast orthologue, Yen1, were independently identified using two distinct experimental approaches: GEN1 was identified by mass spectrometry following extensive fractionation of HeLa cell-free extracts, whereas Yen1 was detected by screening a yeast gene fusion library for nucleases capable of Holliday junction resolution. The eukaryotic Holliday junction resolvases represent a new subclass of the Rad2/XPG family of nucleases. Recombinant GEN1 and Yen1 resolve Holliday junctions by the introduction of symmetrically related cuts across the junction point, to produce nicked duplex products in which the nicks can be readily ligated.

In 1964, Robin Holliday proposed the formation of four-way DNA junctions to account for the types of products that are formed during meiotic recombination in fungi¹. Since that time, the Holliday junction has been invoked as a central intermediate in many models for recombination and the recombinational repair of DNA double-stranded breaks². Because they provide a covalent linkage between chromosomes, their resolution is necessary for chromosome segregation. Enzymes that resolve Holliday junctions by endonucleolytic cleavage have been isolated from bacteriophages T4 and T7, bacteria and archaea³. These proteins are generally small homodimeric nucleases that show both structural and sequence specificities. Resolution occurs by the introduction of symmetrically related nicks in two strands of like polarity, such that the nicked duplex DNA products can be ligated without need for further processing.

The *E. coli* Holliday junction resolvase, RuvC protein, represents the model for the involvement of this class of enzyme in recombination and repair⁴. So far, little is known about eukaryotic Holliday junction resolvases. Holliday junction resolvases have been identified from *S. cerevisiae* (Cce1) and *Schizosaccharomyces pombe* (Ydc2), but these are active only in mitochondria^{5,6}. Mammalian cell-free extracts contain a nuclear Holliday junction resolvase activity (designated ResA)^{7–11}, but its identity remains unknown despite almost two decades of study. This is due to its low abundance, and the fact that mammalian cells (and all eukaryotes) possess alternative means to process Holliday junctions. For example, Holliday junctions can be ‘dissolved’ by the Bloom’s syndrome complex (BLM helicase–topoisomerase III α –RMI1/2)¹², or be cleaved asymmetrically by MUS81–EME1 (asymmetric cleavage produces flapped and gapped duplex products that cannot be ligated without further processing)^{10,13–15}.

Here, we used a two-pronged approach to identify eukaryotic resolvases that fit the classical Holliday junction resolvase model. First, large quantities of cell-free extracts were prepared from human cells and extensively fractionated to enable the identification of ResA by mass spectrometry. Second, a tandem affinity purification (TAP)

fusion library of *S. cerevisiae* genes was screened for proteins that promote Holliday junction resolution. Notably, these parallel approaches led us to identify the same protein: *Homo sapiens* GEN1 and its *S. cerevisiae* orthologue Yen1. We show that GEN1 and Yen1 are highly specific Holliday junction resolvases as defined by the RuvC model.

Identification of the human Holliday junction resolvase

Nuclear extracts prepared from 200 litres of HeLa cells (at 1×10^6 cells ml⁻¹) were fractionated and assayed for Holliday junction resolution activity using the small 5′-³²P-end-labelled synthetic Holliday junction X26, which has a 26-base pair (bp) homologous core through which the junction can migrate. The products of resolution, ³²P-labelled nicked duplex DNA, were visualized by neutral polyacrylamide gel electrophoresis (PAGE) and autoradiography. The purification scheme used a combination of six column matrices that separate proteins according to their size, charge and specific chemical affinities, and was designed to separate ResA away from the MUS81–EME1 complex (Supplementary Fig. 1). Robust Holliday junction resolution activity was present in the peak fractions that eluted from the final monoS column (Fig. 1a, fractions 28–33), but SDS–PAGE analysis showed that these fractions remained quite impure, despite extensive fractionation (Fig. 1b). Indeed, mass spectrometry (MS) revealed that peak fraction 29 contained 75 distinct proteins.

To identify the nuclease responsible for resolution, the proteins in fraction 29 were separated by SDS–PAGE, and the gel was cut into small slices (Fig. 1c, slices a–l). We found that Holliday junction resolution activity could be renatured from a single gel slice (slice e), in which proteins had an expected mass of approximately 60 kDa. MS analyses revealed the identity of 20 proteins in gel slice e, but this was reduced to 12 by subtraction of those also present in other gel slices that lacked resolution activity. One of the remaining candidate proteins, identified by the presence of five peptides by MS analyses, caught our interest as it corresponded to an amino-terminal fragment of an uncharacterized 103 kDa protein (FLJ40869) that

¹Genetic Recombination Laboratory, ²Protein Analysis and Proteomics Laboratory, Cancer Research UK, London Research Institute, Clare Hall Laboratories, South Mimms, Herts EN6 3LD, UK.

*These authors contributed equally to this work.

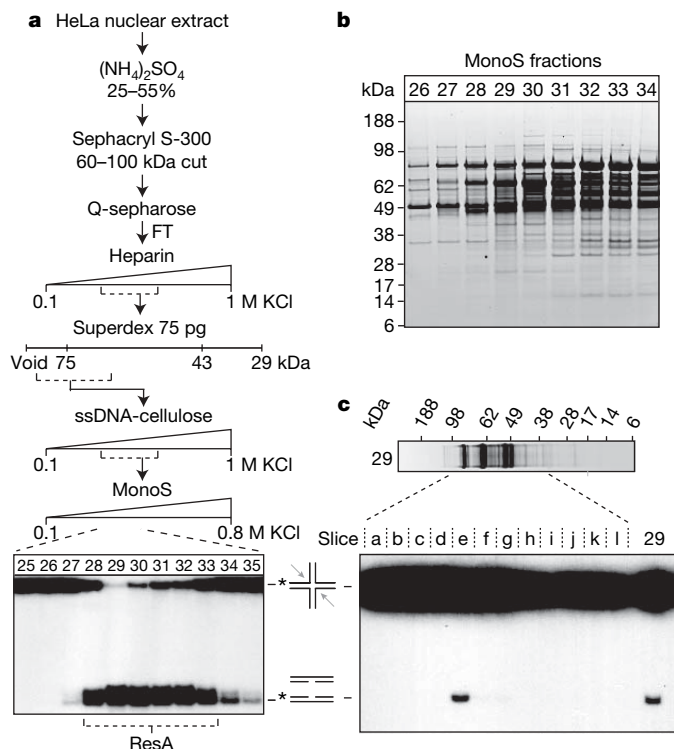


Figure 1 | Identification of human Holliday junction resolvase activity.

a, Purification scheme for the Holliday junction resolvase ResA. Symmetrical cleavage of ^{32}P -labelled Holliday junction X26 (arrows) gives rise to nicked duplex products. Peak fractions 25–35 of the final monoS column are shown. Asterisks indicate ^{32}P -labels. FT, flow through. **b**, SDS-PAGE of fractions 26–34 from monoS, stained with SYPRO Ruby. **c**, Fraction 29 from **b** was cut into slices (a–l), and proteins were renatured and analysed for Holliday junction resolution using ^{32}P -labelled Holliday junction X0. Control lane (right): fraction 29 diluted 50-fold.

contains a Rad2/XPG nuclease domain (Supplementary Fig. 2). This human protein was identified as GEN1, which is encoded by a gene at chromosome location 2p24.2.

Identification of the yeast Holliday junction resolvase

In parallel with the above approach, a *S. cerevisiae* TAP fusion library¹⁶ was screened for Holliday junction resolution activities. The library contains 4,247 open reading frames tagged and expressed from their genomic locus. Strains harbouring TAP fusion proteins were individually grown in 100 ml cultures and TAP fusion proteins were immunoprecipitated from their lysates and tested for resolution using ^{32}P -labelled Holliday junction X26. The screen was biased towards the low expression categories (50–1,500 molecules per cell), based on our prediction that a Holliday junction resolvase would not be an abundant protein. We also excluded proteins that have well established and unrelated functions (for example, membrane proteins). In total we screened 1,100 strains.

Representative results are shown in Fig. 2. The screening procedure was optimized with a yeast strain expressing TAP-tagged Cce1 (present at approximately 50 molecules per cell), which was used as an internal control in all analyses. Cce1 was the first TAP fusion strain in the library (strain 76) to give a positive result in the Holliday junction resolution assay (Fig. 2a), which generally gave clear negatives for almost all other TAP fusions. We next identified a Holliday junction processing activity associated with Mms4 (strain 152; Mms4 is known as Eme1 in *S. pombe* and EME1 in mammalian cells and is a component of the Mus81–Mms4 heterodimeric nuclease) (Fig. 2b), Mus81 itself (strain 520) and Slx4 (strain 582), which is known to form a complex with Slx1 (data not shown). Although Slx1–Slx4, a flap nuclease, has detectable activity with Holliday junctions (like

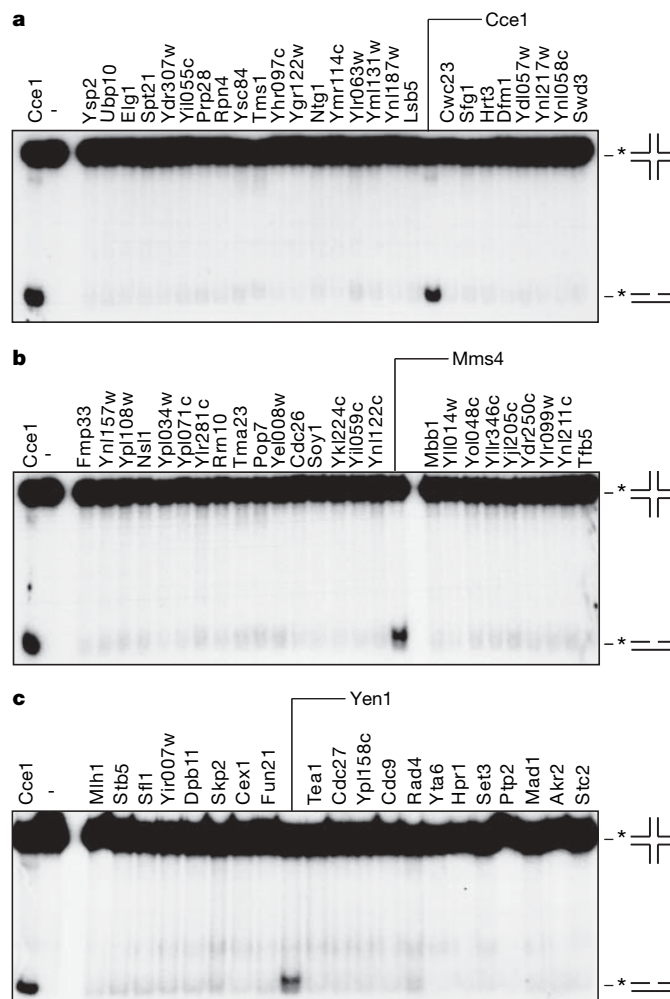


Figure 2 | Yeast screen for Holliday junction resolution activities.

a–c, Lysates were prepared from 1,100 yeast cultures containing *S. cerevisiae* TAP fusion proteins. The TAP fusion proteins were immunoprecipitated using IgG sepharose and analysed for their ability to promote the resolution of ^{32}P -labelled Holliday junction X26. Three representative panels of the high-throughput screen are shown. Each lane represents one candidate protein, indicated by protein name or open reading frame code.

Mus81 the cuts are asymmetric) they are not thought to represent its physiological substrate¹⁷.

We also isolated novel candidate Holliday junction resolvases, of which a protein known as Yen1 (strain 570) looked the most promising (Fig. 2c). Yen1 is encoded by open reading frame YER041W, and its function is unknown (*yen1* mutants are viable). During further study, Yen1–TAP displayed consistent Holliday junction cleavage activity at levels comparable to Cce1 and Mus81–Mms4 (Supplementary Fig. 3). Further isolates included Cef1, Mgs1, Apn2, Rad16 and Slx8, but subsequent tests failed to confirm these proteins as capable of Holliday junction resolution (Supplementary Fig. 3). This left us with Yen1 as the only viable candidate for a novel Holliday junction resolution activity. Notably, Yen1 is the *S. cerevisiae* orthologue of human GEN1 (Supplementary Fig. 4), the Holliday junction resolution activity identified from HeLa cells. Thus, two completely different approaches to identify Holliday junction resolvases led us to the same protein. We therefore focused our efforts on Yen1 and GEN1.

Holliday junction cleavage by Yen1 and GEN1

Yen1 and GEN1 are members of the Rad2/XPG family of structure-specific nucleases that are characterized by the presence of three distinct sequence motifs: the N-terminal (N) and internal (I) XPG

nuclease domains, together with a helix–hairpin–helix domain (Supplementary Fig. 5)¹⁸. The Rad2/XPG family contains several well-characterized nucleases, but in contrast to Yen1–TAP, we did not observe specific Holliday junction resolution by other family members such as Rad2–TAP, Rad27–TAP or Exo1–TAP (Supplementary Fig. 6). These results suggest that Yen1 and GEN1 may represent a subclass of the Rad2/XPG family that has evolved to resolve Holliday junctions.

To determine whether Holliday junction cleavage occurred by symmetric (RuvC-like resolution) or asymmetric (like Mus81–Eme1) cutting, we prepared vectors that expressed recombinant forms of full-length Yen1 and GEN1. First, *YEN1* was placed under the control of the *GAL1* promoter and Yen1–Flag was overexpressed in yeast strain W303. When cell-free extracts were assayed for Holliday junction resolution activity using a Holliday junction with a fixed branch point (Holliday junction X0), we found that Yen1 expression resulted in efficient Holliday junction resolution (Fig. 3a, lane 3). In contrast, a mutant form of Yen1, created by mutating two conserved glutamic acid residues (residues 193 and 195) to alanine (Supplementary Fig. 4), or extracts prepared from cells carrying the empty vector, failed to resolve X0 (Fig. 3a, lanes 2 and 4). Analogous mutations abolish¹⁹ or severely reduce²⁰ the nuclease activities of other Rad2/XPG family proteins. Second, human *GEN1* was transiently transfected into human 293T cells, and GEN1–Flag protein was partially purified by Flag affinity chromatography. We found that GEN1–Flag also resolved Holliday junction X0 efficiently (Fig. 3b, lane 3).

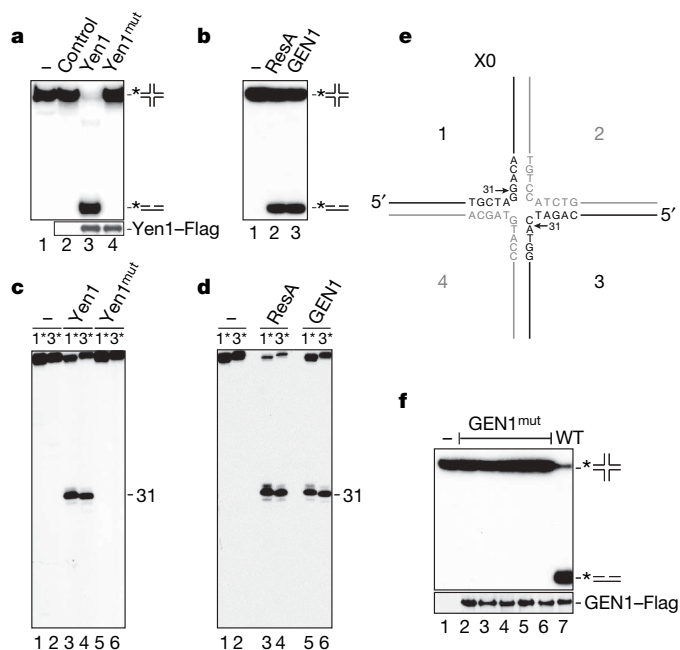


Figure 3 | Resolution of Holliday junctions by recombinant Yen1 and GEN1. **a**, ³²P-labelled Holliday junction X0 was incubated with cell-free extracts from yeast overexpressing Flag-tagged Yen1, or a catalytically inactive Yen1 (E193A/E195A) mutant (Yen1^{mut}), and the products were analysed by neutral PAGE. Control: extracts from cells transformed with empty expression vector. Yen1–Flag and Yen1^{mut}–Flag were detected by western blotting using anti-Flag antibody. **b**, As **a**, but affinity purified GEN1–Flag, or ResA, was used. **c**, **d**, Holliday junction X0, 5′-³²P-end-labelled in strand 1 or 3, was treated as in **a** and **b**, and products were analysed by denaturing PAGE. Asterisks indicate the strand with the radioactive label. **e**, The cleavage sites. **f**, Wild-type and mutant derivatives of GEN1–Flag were assayed using Holliday junction X0. Lane 1, control lane; lane 2, GEN1(D30A); lane 3, GEN1(E134A); lane 4, GEN1(E136A); lane 5, GEN1(D157A); lane 6, GEN1(E134A/E136A); lane 7, wild-type GEN1. In the bottom panel, GEN1–Flag proteins were detected by western blotting.

Using Holliday junction X0, 5′-³²P-end-labelled in either strand 1 or 3, we observed that both Yen1 and GEN1 cut the immobile junction X0 at a unique site located one nucleotide to the 3′ side of the junction point. The nicks were introduced with perfect symmetry to produce fragments that were 31 nucleotides in length as determined by denaturing PAGE (Fig. 3c, lanes 3 and 4, and Fig. 3d, lanes 5 and 6). The cleavage pattern is summarized in Fig. 3e. Recombinant GEN1–Flag produced identical cuts to those produced by ResA purified from HeLa cells (Fig. 3d, compare lanes 3 and 4 with 5 and 6).

Yen1 and GEN1 also produced perfectly symmetric cleavage patterns within the region of homology in Holliday junction X26, with cleavage occurring at two preferred sites within the 26-bp homologous core (Supplementary Figs 7 and 8). Thus, like RuvC, the eukaryotic resolvases show some sequence preferences. The cleavage pattern produced by GEN1 was again identical to that produced by ResA (Supplementary Fig. 8), confirming that GEN1 is equivalent to ResA from HeLa cells. Importantly, the introduction of site-directed mutations in the nuclease domains of GEN1–Flag resulted in a complete loss of Holliday junction resolution activity (Fig. 3f).

Symmetrical cleavage of Holliday junctions by RuvC results in the formation of nicked duplex products in which the nicks can be ligated without need for further processing. To determine whether the eukaryotic Holliday junction resolvase gives rise to ligatable products, we constructed a derivative of the migratable Holliday junction X26 in which the arms were of different length (the junction was 5′-³²P-end-labelled in the 53-nucleotide-long strand 2S) such that nick ligation could be assayed by denaturing PAGE as the formation of a 60-nucleotide-long ³²P-labelled strand (shown schematically in Fig. 4a). When resolved by either RuvC or GEN1–Flag (Fig. 4b, lanes 3 and 5), and treated with T4 DNA ligase, equivalent levels of nick religation were observed (lanes 4 and 6). Thus, both RuvC and GEN1 leave ligatable nicks after symmetrical cleavage.

Substrate specificity

Given that Rad2/XPG family members are structure-specific nucleases, we next determined the specificity of GEN1–Flag. For this, the protein was incubated with equimolar amounts of a variety of substrates including splayed-arm DNA, 3′- and 5′-flap DNAs, and model replication fork DNA. GEN1–Flag showed a strong preference for the Holliday junction (Fig. 4c). Time course experiments indicated that the initial rate of Holliday junction cleavage was 7- and 20-fold greater than that observed with 5′-flap or replication fork substrates, respectively (Fig. 4d). GEN1 failed to cut splayed-arm or 3′-flap substrates.

To determine whether GEN1 alone was capable of Holliday junction resolution, or whether cleavage required the presence of other co-purifying factors, we took advantage of the fact that the N-terminal region of GEN1 contains the Rad2/XPG nuclease domains and purified GEN1(1–527) to near homogeneity from *E. coli* (Fig. 4e). Recombinant GEN1(1–527) was an active Holliday junction resolvase (Fig. 4f), and showed the same structure specificity as affinity-purified human GEN1–Flag protein. This specificity for Holliday junctions, coupled with the observed symmetrical cleavage mechanism, shows that Yen1 and GEN1 are bona fide Holliday junction resolvases similar to RuvC.

Discussion

When the *E. coli* Holliday junction resolvase, RuvC protein, was discovered in 1991 (refs 21, 22), and analogous activities were detected in mammalian cell-free extracts⁷, it seemed that the mechanisms of Holliday junction resolution were conserved throughout all kingdoms of life. However, this concept was thrown into question by the description of alternative ways to process Holliday junctions involving either topoisomerase-mediated junction dissolution¹², or asymmetric cleavage by MUS81–EME1 flap endonucleases^{13,23}. The identification of the novel eukaryotic GEN1/Yen1 Holliday junction resolvases described here, and analysis of their mechanism of resolution, suggests that the RuvC model is indeed universal, while also

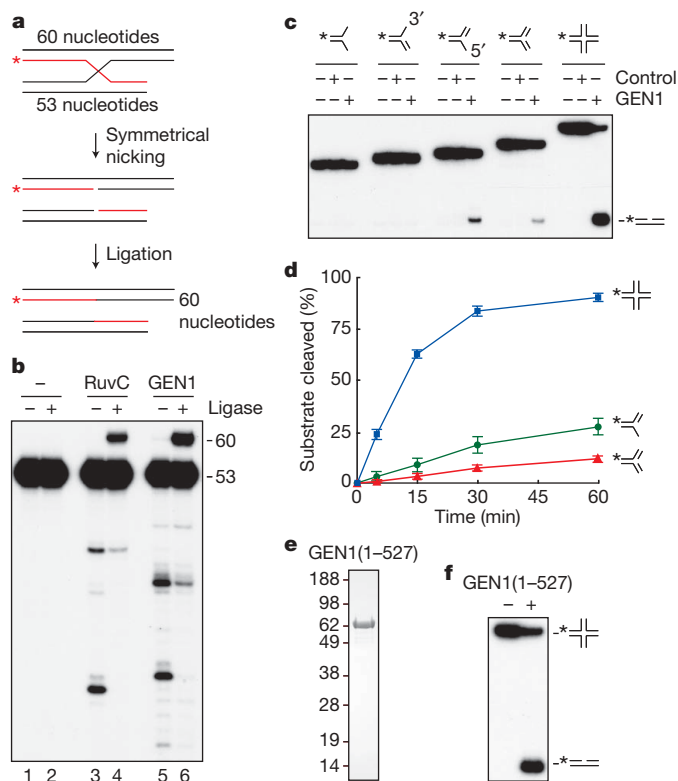


Figure 4 | Specificity of GEN1 Holliday junction resolvase. **a**, Assay for nick ligation. Holliday junction X26-S, 32 P-end-labelled in strand 2S (red: 53-nucleotides long), was treated with GEN1-Flag (or RuvC) and the resolution products were incubated with T4 DNA ligase. Symmetrical cleavage followed by ligation produces a 60-nucleotide-long 32 P-labelled strand. **b**, Ligation of resolution products generated by either RuvC or GEN1-Flag, determined by denaturing PAGE. **c**, Substrate specificity of GEN1-Flag. The indicated substrates (from left to right: splayed arm, 3'-flap, 5'-flap, replication fork, and Holliday junction), 32 P-end-labelled on a common strand, were incubated with GEN1-Flag or a mock protein control. Products were analysed by neutral PAGE. **d**, Time course of reactions similar to those of **c**. Product formation was quantified by phosphorimaging and expressed as a percentage of total DNA (\pm s.d., $n = 3$). **e**, SDS-PAGE gel showing Coomassie-blue-stained His-tagged GEN1(1-527) (11 μ g) purified from *E. coli*. **f**, Holliday junction resolution by His-tagged GEN1(1-527) (60 nM).

highlighting the diversification of Holliday junction processing mechanisms in eukaryotes.

GEN1 and Yen1 promote Holliday junction resolution by a symmetrical cleavage mechanism analogous to that shown by RuvC, Cce1, Hjc and the bacteriophage resolvases. However, they do not show any sequence or structure conservation with these proteins, but instead belong to the Rad2/XPG family of nucleases¹⁸. The presence of the eukaryotic Holliday junction resolvases in the Rad2/XPG nuclease family emphasizes the wide-ranging functions of this group of proteins. The phylogenetic tree (Supplementary Fig. 9) shows the family divided into four subclasses, consistent with their cellular functions: class I, characterized by XPG (*H. sapiens*) and Rad2 (*S. cerevisiae*), promote endonucleolytic incision at the 3' side of bubble structures formed during nucleotide excision repair²⁴; class II, characterized by FEN1 (*H. sapiens*) and Rad27 (*S. cerevisiae*), use their 5'-flap endonuclease activities for Okazaki fragment processing during DNA replication and also act in a variety of specialized DNA repair reactions that require the trimming of 5' flaps²⁵; class III comprise the EXO1-related proteins, which exhibit 5'-3' exonuclease and weak 5'-flap endonuclease activities, that are required for DNA replication, DNA repair and meiotic recombination²⁶; and the results described here indicate that class IV comprise the eukaryotic Holliday junction resolvases.

The yeast *YEN1* and human *GEN1* genes encode proteins of 87 kDa and 103 kDa, respectively. However, our studies show that a 60 kDa N-terminal fragment of GEN1 (GEN1(1-527)) purified from *E. coli* promotes efficient Holliday junction resolution, as does the truncated form of GEN1 purified from HeLa extracts. It is therefore unclear whether the protein is functionally active as a full-length protein. One possibility is that the resolvase is activated by proteolytic cleavage or by interaction with other factors. In this regard, it is interesting that extracts prepared from cell lines defective in the RAD51C or XRCC3 recombination/repair proteins show reduced levels of Holliday junction resolution activity¹¹. However, although RAD51C and XRCC3 have been reported to associate with ResA, we did not detect the presence of these proteins in the most purified ResA/GEN1 fraction from HeLa cells, or in the affinity-purified GEN1-Flag fractions (Supplementary Fig. 10). These data indicate that no strong physical link exists between RAD51C-XRCC3 and GEN1, and that these proteins are not directly required for Holliday junction resolution by purified GEN1. Further work will be required to define the relationships between RAD51C-XRCC3 and GEN1.

Eukaryotic cells process recombination intermediates in several different ways, and the outcome is likely to be dependent on pathway choice (Supplementary Fig. 11). In mammalian cells Holliday junction intermediates may be dissolved by BLM-topoisomerase III α -RMI1/2 complex, to produce non-crossovers (in which there is no exchange of flanking markers), or be cleaved by MUS81-EME1 complex to produce crossovers with an exchange of flanking markers. The efficient Holliday junction resolution activities described here are indicative of a third pathway, Holliday junction resolution by Yen1/GEN1, which by analogy with other resolvases would be expected to give rise to crossovers and non-crossovers with equal efficiency, depending on the orientation of Holliday junction cleavage.

S. pombe mus81 and *eme1* mutants have pronounced meiotic recombination defects consistent with defects in the processing of recombination intermediates^{23,27}. Indeed, Mus81-Eme1 seems to provide the primary means to generate crossovers in this organism, although its actions may not be restricted to Holliday junctions²⁸. In contrast, *S. cerevisiae mus81* and *mms4* mutants have only mild defects in crossover formation²⁹, and *Mus81*^{-/-} mice fail to show any obvious meiotic defects^{30,31}. The heightened dependence that *S. pombe* shows for Mus81-Eme1 may be due to the fact that *YEN1* is absent from the *S. pombe* genome, whereas it is present in *S. cerevisiae* and other eukaryotes. Consistent with the concept of alternative mechanisms by which joint molecules can be resolved, it was recently shown that crossover and non-crossover products still form at high levels in *S. cerevisiae sgs1* (an orthologue of BLM) *mms4* double mutants^{32,33}. Preliminary studies in *S. cerevisiae* indicate that *yen1 mms4* double mutants fail to complete meiosis and contain unresolved interhomologue joint molecules (L. Jessop, E. Kolar, J. Haber and M. Lichten, personal communication), suggestive of functional redundancy between Mus81-Mms4 and Yen1.

Previously, it was shown that weak ssDNA and 5'-flap/fork endonuclease activities were associated with GEN1 orthologues from rice^{34,35} and insects^{36,37}, and that RNA-interference-mediated gene silencing of rice *GEN1* resulted in male sterility³⁵. Additionally, the messenger RNA levels of a rice GEN1 orthologue were found to be induced by DNA-damaging agents such as ultraviolet light, methyl methanesulphonate or H₂O₂, consistent with a role in DNA repair³⁴. Genome-wide sequencing of all coding regions in tumours derived from a cohort of breast cancer patients revealed deletions/frameshift mutations in *GEN1* (FLJ40869) in two individuals³⁸. As the mutations found within *GEN1* have a high probability of being pathogenic, rather than passenger mutations, *GEN1* was classified as a cancer gene. These mutations were 785delC and 824-827_{del}GTAA, which would give rise to GEN1 truncations without Holliday junction resolution activity. Thus, *GEN1* joins the growing list of genes involved in recombinational repair, such as *BRCA1*,

BRCA2, *FANCF* (also called *BACH1*) and *FANCN* (*PALB2*), that are linked with tumorigenesis.

METHODS SUMMARY

DNA substrates, 5'-³²P-end-labelled on oligonucleotides X26-2 or X0-1, were made by annealing oligonucleotides (Supplementary Table 1) and purified by PAGE³⁹. Yeast TAP fusion library¹⁶ was purchased from Open Biosystems, and screened as described³⁹. Yen1-Flag was expressed in yeast and soluble lysates were analysed for Holliday junction resolution activity. Human GEN1-Flag was purchased from Origene as a pCMV-Entry expression vector, transfected into 293T cells, and affinity purified using anti-Flag M2 affinity gel (Sigma). A mock protein control was made from untransfected 293T cells using the same procedure. Mutant derivatives of GEN1-Flag were made by site-directed mutagenesis. Carboxy-terminal His-tagged GEN1 (1–527) was purified after overexpression in *E. coli*, using HisTrap, heparin, ssDNA and monoS columns. Holliday junction resolution and substrate specificity assays were performed under similar conditions. For the ligation experiment, cleavage reactions were supplemented with 400 U T4 DNA ligase (NEB) and incubation continued at room temperature for 1 h. ³²P-labelled products were analysed by 10% neutral PAGE, or by electrophoresis through a 12% denaturing gel containing 7 M urea.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 4 August; accepted 26 September 2008.

- Holliday, R. A mechanism for gene conversion in fungi. *Genet. Res. Camb.* **5**, 282–304 (1964).
- Paques, F. & Haber, J. E. Multiple pathways of recombination induced by double-strand breaks in *Saccharomyces cerevisiae*. *Microbiol. Mol. Biol. Rev.* **63**, 349–404 (1999).
- Lilley, D. M. J. & White, M. F. The junction-resolving enzymes. *Nature Rev. Mol. Cell Biol.* **2**, 433–443 (2001).
- West, S. C. Processing of recombination intermediates by the RuvABC proteins. *Annu. Rev. Genet.* **31**, 213–244 (1997).
- Kleff, S., Kemper, B. & Sternglanz, R. Identification and characterization of yeast mutants and the gene for a cruciform cutting endonuclease. *EMBO J.* **11**, 699–704 (1992).
- Whitby, M. C. & Dixon, J. Substrate specificity of the SpCCE1 Holliday junction resolvase of *Schizosaccharomyces pombe*. *J. Biol. Chem.* **273**, 35063–35073 (1998).
- Elborough, K. M. & West, S. C. Resolution of synthetic Holliday junctions in DNA by an endonuclease activity from calf thymus. *EMBO J.* **9**, 2931–2936 (1990).
- Hyde, H., Davies, A. A., Benson, F. E. & West, S. C. Resolution of recombination intermediates by a mammalian endonuclease activity functionally analogous to *Escherichia coli* RuvC resolvase. *J. Biol. Chem.* **269**, 5202–5209 (1994).
- Constantinou, A., Davies, A. A. & West, S. C. Branch migration and Holliday junction resolution catalyzed by activities from mammalian cells. *Cell* **104**, 259–268 (2001).
- Constantinou, A., Chen, X.-B., McGowan, C. H. & West, S. C. Holliday junction resolution in human cells: Two junction endonucleases with distinct substrate specificities. *EMBO J.* **21**, 5577–5585 (2002).
- Liu, Y., Masson, J.-Y., Shah, R., O'Regan, P. & West, S. C. RAD51C is required for Holliday junction processing in mammalian cells. *Science* **303**, 243–246 (2004).
- Wu, L. & Hickson, I. D. The Bloom's syndrome helicase suppresses crossing over during homologous recombination. *Nature* **426**, 870–874 (2003).
- Chen, X. B. *et al.* Human MUS81-associated endonuclease cleaves Holliday junctions *in vitro*. *Mol. Cell* **8**, 1117–1127 (2001).
- Ciccia, A., Constantinou, A. & West, S. C. Identification and characterization of the human MUS81-EME1 endonuclease. *J. Biol. Chem.* **278**, 25172–25178 (2003).
- Taylor, E. R. & McGowan, C. H. Cleavage mechanism of human MUS81-EME1 acting on Holliday-junction structures. *Proc. Natl Acad. Sci. USA* **105**, 3757–3762 (2008).
- Ghaemmaghami, S. *et al.* Global analysis of protein expression in yeast. *Nature* **425**, 737–741 (2003).
- Fricke, W. M. & Brill, S. J. Slx1-Slx4 is a second structure-specific endonuclease functionally redundant with Sgs1-Top3. *Genes Dev.* **17**, 1768–1778 (2003).
- Lieber, M. R. The FEN-1 family of structure-specific nucleases in eukaryotic DNA replication, recombination and repair. *Bioessays* **19**, 233–240 (1997).
- Shen, B., Nolan, J. P., Sklar, L. A. & Park, M. S. Functional analysis of point mutations in human flap endonuclease-I active site. *Nucleic Acids Res.* **25**, 3332–3338 (1997).
- Tran, P. T., Erdeniz, N., Dudley, S. & Liskay, R. M. Characterization of nuclease-dependent functions of Exo1p in *Saccharomyces cerevisiae*. *DNA Repair* **1**, 895–912 (2002).
- Dunderdale, H. J. *et al.* Formation and resolution of recombination intermediates by *E. coli* RecA and RuvC proteins. *Nature* **354**, 506–510 (1991).
- Iwasaki, H., Takahagi, M., Shiba, T., Nakata, A. & Shinagawa, H. *Escherichia coli* RuvC protein is an endonuclease that resolves the Holliday structure. *EMBO J.* **10**, 4381–4389 (1991).
- Boddy, M. N. *et al.* Mus81-Eme1 are essential components of a Holliday junction resolvase. *Cell* **107**, 537–548 (2001).
- O'Donovan, A., Davies, A. A., Moggs, J. G., West, S. C. & Wood, R. D. XPG endonuclease makes the 3' incision in human DNA nucleotide excision repair. *Nature* **371**, 432–435 (1994).
- Liu, Y., Kao, H. I. & Bambara, R. A. Flap endonuclease I: a central component of DNA metabolism. *Annu. Rev. Biochem.* **73**, 589–615 (2004).
- Tran, P. T., Erdeniz, N., Symington, L. S. & Liskay, R. M. EXO1 - A multi-tasking eukaryotic nuclease. *DNA Repair* **3**, 1549–1559 (2004).
- Smith, G. R., Boddy, M. N., Shanahan, P. & Russell, P. Fission yeast Mus81-Eme1 Holliday junction resolvase is required for meiotic crossing over but not for gene conversion. *Genetics* **165**, 2289–2293 (2003).
- Osman, F., Dixon, J., Doe, C. L. & Whitby, M. C. Generating crossovers by resolution of nicked Holliday junctions: A role of Mus81-Eme1 in meiosis. *Mol. Cell* **12**, 761–774 (2003).
- de los Santos, T. *et al.* The Mus81-Mms4 endonuclease acts independently of double-Holliday junction resolution to promote a distinct subset of crossovers during meiosis in budding yeast. *Genetics* **164**, 81–94 (2003).
- Dendouga, N. *et al.* Disruption of murine Mus81 increases genomic instability and DNA damage sensitivity but does not promote tumorigenesis. *Mol. Cell Biol.* **25**, 7569–7579 (2005).
- McPherson, J. P. *et al.* Involvement of mammalian Mus81 in genome integrity and tumor suppression. *Science* **304**, 1822–1826 (2004).
- Oh, S. D., Lao, J. P., Taylor, A. F., Smith, G. R. & Hunter, N. RecQ helicase, Sgs1, and XPF family endonuclease, Mus81-Mms4, resolve aberrant joint molecules during meiotic recombination. *Mol. Cell* **31**, 324–336 (2008).
- Jessop, L. & Lichten, M. Mus81/Mms4 endonuclease and Sgs1 helicase collaborate to ensure proper recombination intermediate metabolism during meiosis. *Mol. Cell* **31**, 313–323 (2008).
- Furukawa, F. *et al.* OsSEND-1: a new Rad2 nuclease family member in higher plants. *Plant Mol. Biol.* **51**, 59–70 (2003).
- Moritoh, S. *et al.* RNAi-mediated silencing of OsGEN-L (OsGEN-like), a new member of the Rad2/XPG nuclease family, causes male sterility by defect of microspore development in rice. *Plant Cell Physiol.* **46**, 699–715 (2005).
- Ishikawa, G. *et al.* DmGEN, a novel Rad2 family endo-exonuclease from *Drosophila melanogaster*. *Nucleic Acids Res.* **32**, 6251–6259 (2004).
- Kanai, Y. *et al.* DmGEN shows a flap endonuclease activity, cleaving the blocked-flap structure and model replication fork. *FEBS J.* **274**, 3914–3927 (2007).
- Wood, L. D. *et al.* The genomic landscapes of human breast and colorectal cancers. *Science* **318**, 1108–1113 (2007).
- Rass, U. & West, S. C. Synthetic junctions as tools to identify and characterise Holliday junction resolvases. *Methods Enzymol.* **408**, 485–501 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank M. Lichten and J. Haber for communication of unpublished work, J.-Y. Bleuyard, A. Deans and other members of our laboratory, past and present, for their input and discussions. We acknowledge the contributions of K. Elborough, H. Hyde, B. Kysela, A. Davies, A. Constantinou and Y. Liu for early studies that underpin this work, which was supported by Cancer Research UK, the Louis-Jeantet Foundation and the EU DNA Repair Consortium. S.C.Y.I. was supported by the Croucher Foundation. M.G.B. is on leave from the University of Santiago de Compostela, and supported by fellowships from the Xunta de Galicia and the Spanish Ministry of Education.

Author Contributions S.C.Y.I. identified ResA as GEN1 and characterized the recombinant protein. U.R. designed and carried out the yeast screen to identify Yen1. M.G.B. expressed and characterized recombinant Yen1. H.R.F. carried out mass spectrometric analyses and together with J.M.S. analysed the MS data. S.C.W. helped with experimental design and wrote the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to S.C.W. (stephen.west@cancer.org.uk).

METHODS

Identification of human Holliday junction resolvase. HeLa S3 cells (2001 at 1×10^6 cells ml^{-1}) were grown¹⁰, harvested and washed three times with iced PBS. The cell pellet was re-suspended in 2 vol hypotonic lysis buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA, 2 mM dithiothreitol (DTT)), left on ice for 1 h and lysed using an 'A' pestle (25 strokes) in the presence of complete EDTA-free protease inhibitor cocktail (Roche). After centrifugation (1 h at 2,500g), the nuclear pellet was resuspended in 2 vol high-salt extraction buffer (10 mM Tris-HCl pH 8.0, 300 mM KCl, 1 mM EDTA, 2 mM DTT) and incubated on ice for 1 h before being centrifuged again (1 h at 35,000 r.p.m., Beckman 45Ti rotor). Solid $(\text{NH}_4)_2\text{SO}_4$ (25%; 134 g l^{-1}) was added to the supernatant and dissolved by stirring on ice. After 30 min insoluble material was removed by centrifugation (30 min at 12,000g). The $(\text{NH}_4)_2\text{SO}_4$ concentration was raised to 55% (an additional 179 g l^{-1}), and precipitated proteins were recovered by centrifugation and re-suspended in 32 ml buffer A (50 mM $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$ pH 6.8, 10% glycerol, 1 mM EDTA, 1 mM DTT and 0.01% NP40) containing 250 mM KCl. After dialysis against the same buffer, proteins were loaded onto a 320 ml sephacryl S-300 column in four separate runs. Fractions (112 ml) with resolution activity were pooled, dialysed against buffer B (50 mM $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$ pH 7.4, 10% glycerol, 1 mM EDTA, 1 mM DTT and 0.01% NP40) containing 100 mM KCl flowed through a 1 ml HiTrap Q sepharose (GE healthcare) and loaded onto a 5 ml HiTrap heparin column (GE healthcare). After washing, proteins were eluted with a 75 ml gradient to 1 M KCl. Resolvase fractions were loaded onto a 120 ml superdex 75 column, pre-equilibrated with buffer A plus 250 mM KCl, eluted and dialysed against buffer B plus 100 mM KCl before loading onto a 300 μl ssDNA cellulose column (Amersham). The column was eluted with a 4.5 ml gradient to 1 M KCl, and fractions were dialysed against buffer B plus 100 mM KCl before loading onto a monoSP 1.6/5 column (GE healthcare). Resolution activity was eluted in a 1 ml gradient to 800 mM KCl. Assays were carried out using Holliday junction X26 in the presence of nonspecific competitor DNA (10 $\text{ng } \mu\text{l}^{-1}$ poly(dI-dC)).

An aliquot of the peak ResA fraction was heated for 10 min at 72 °C in NuPAGE LDS sample buffer and run on a Novex 4–12% Bis-Tris PAGE gel in MES SDS running buffer (Invitrogen). The gel was washed twice in 20 ml elution-renaturation buffer (60 mM $\text{K}_2\text{HPO}_4/\text{KH}_2\text{PO}_4$ pH 7.4, 1 mM DTT, 100 $\mu\text{g ml}^{-1}$ BSA, 50 mM KCl, 1% Triton X-100), and cut into small slices. Each slice was homogenized in 3 vol of elution-renaturation buffer and incubated at 4 °C for 5 h with gentle mixing. Supernatants containing eluted proteins were analysed for Holliday junction resolution activity by overnight incubation at 37 °C with ^{32}P -labelled Holliday junction X0 (~ 1 nM) and 5 mM $\text{Mg}(\text{OAc})_2$.

Peptides for analysis were generated by *in situ* tryptic digestion of protein/gel bands. LC-ESI/MS/MS analysis of the peptides was carried out on a SYNAPT HDMS mass spectrometer (Waters) and the data searched against a concatenated, non-redundant protein database (UniProt 13.6), using the Mascot search engine (Matrix Science).

Resolution assays. Holliday junction resolution reactions (10 μl) contained ^{32}P -labelled DNA substrates (~ 1 nM) and 0.5 μl aliquots of the indicated mammalian fractions in phosphate buffer (60 mM $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ pH 7.4, 1 mM DTT, 100 $\mu\text{g ml}^{-1}$ BSA, 2 mM ATP) supplemented with 5 mM

$\text{Mg}(\text{OAc})_2$. For Yen1, reactions contained ^{32}P -labelled DNA (~ 1 nM) and protein extract (5 μg) in 40 mM Tris-HCl pH 7.5, 1 mM MgCl_2 and 10 mM NaCl. Reactions with RuvC (100 nM), purified as described⁴⁰, were carried out in phosphate buffer supplemented with 15 mM $\text{Mg}(\text{OAc})_2$. Incubation was for 30 min at 37 °C (ResA/GEN1) or 30 °C (Yen1), and the DNA products were deproteinized for 15 min at 37 °C using 2 mg ml^{-1} proteinase K and 0.4% SDS.

Preparation of Yen1–Flag. The *YEN1* gene was amplified from W303 yeast genomic DNA, a C-terminal Flag-tag was added, and cloned into pYES-DEST52 containing a *GAL1* promoter for overexpression in yeast. The coding sequence of *YEN1* differed from the *S. cerevisiae* database with a C175G change that resulted in alanine at amino acid 59. This sequence was found in two independent clones of W303 and also in strain BY4741, and is displayed in the NCBI website (Locus P40028). Yen1–Flag was overexpressed in pYEN1–Flag transformed W303 cells, and after 10 h soluble extracts were made by mechanical disruption followed by centrifugation.

Purification of GEN1(1–527). The GEN1(1–527) coding sequence was transferred to the Gateway destination vector pET-DEST42 for overexpression in *E. coli* BL21(DE3)-RIL (Stratagene). GEN1(1–527)-His expression was induced using 100 μM IPTG, and after 7 h cells were harvested and lysed. GEN1(1–527) was purified using a HisTrap FF Crude column, a HiTrap heparin column, a ssDNA cellulose column and finally a monoSP 1.6/5 column (all GE healthcare).

Sequence analysis. Domains present in the Rad2/XPG family members were compared using the SMART database⁴¹. Alignments were carried out using Jaligier (<http://jaligier.sourceforge.net>), using a BLOSUM90 matrix, a gap open value of 10 and a gap extension value of 0.5. The phylogenetic tree was constructed using one or two representatives from each of the classes of the Rad2/XPG family in *Arabidopsis thaliana*, *S. cerevisiae*, *Caenorhabditis elegans*, *Drosophila melanogaster* and *H. sapiens* using sequences from the NCBI database (<http://www.ncbi.nlm.nih.gov>) or WormBase (<http://www.wormbase.org>). Accession numbers of the sequences are: *S. cerevisiae*: NP_011774 (Rad2), NP_012809 (Rad27), NP_014676 (Exo1), NP_010549 (Din7), NP_010959 (Yen1); *A. thaliana*: NP_566830 (XPG), NP_850877 (FEN1), Q8L6Z7 (EXO1), Q9LPD2 (GEN1); *C. elegans*: NP_491891 (XPG), WP:CE28239 (homologous to FEN1), WP:CE28728 (homologous to EXO1), WP:CE01401 (homologous to GEN1); *D. melanogaster*: AAD50779 (XPG), NP_523765 (FEN1), Q24558 (EXO1/tosca), NP_647943 (GEN1); *H. sapiens*: NP_000114 (XPG), NP_004102 (FEN1), NP_003677 (EXO1a), NP_872431 (GEN1). Sequences were aligned and analysed with ClustalX 2.0.9 (ref. 42) using default parameters. A neighbour-joining tree was generated and bootstrapped by 1,000 replications. The tree was visualized with Treeview⁴³.

40. Eggleston, A. K., Mitchell, A. H. & West, S. C. *In vitro* reconstitution of the late steps of genetic recombination in *E. coli*. *Cell* **89**, 607–617 (1997).

41. Letunic, I. et al. SMART 5: domains in the context of genomes and networks. *Nucleic Acids Res.* **34**, D257–D260 (2006).

42. Larkin, M. A. et al. Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).

43. Page, R. D. M. TreeView: An application to display phylogenetic trees on personal computers. *Bioinformatics* **12**, 357–358 (1996).

LETTERS

An excess of cosmic ray electrons at energies of 300–800 GeV

J. Chang^{1,2}, J. H. Adams Jr³, H. S. Ahn⁴, G. L. Bashindzhagyan⁵, M. Christl³, O. Ganel⁴, T. G. Guzik⁶, J. Isbert⁶, K. C. Kim⁴, E. N. Kuznetsov⁵, M. I. Panasyuk⁵, A. D. Panov⁵, W. K. H. Schmidt², E. S. Seo⁴, N. V. Sokolskaya⁵, J. W. Watts³, J. P. Wefel⁶, J. Wu⁴ & V. I. Zatsepin⁵

Galactic cosmic rays consist of protons, electrons and ions, most of which are believed to be accelerated to relativistic speeds in supernova remnants^{1–3}. All components of the cosmic rays show an intensity that decreases as a power law with increasing energy (for example as $E^{-2.7}$). Electrons in particular lose energy rapidly through synchrotron and inverse Compton processes, resulting in a relatively short lifetime (about 10^5 years) and a rapidly falling intensity, which raises the possibility of seeing the contribution from individual nearby sources (less than one kiloparsec away)⁴. Here we report an excess of galactic cosmic-ray electrons at energies of ~300–800 GeV, which indicates a nearby source of energetic electrons. Such a source could be an unseen astrophysical object (such as a pulsar⁵ or micro-quasar⁶) that accelerates electrons to those energies, or the electrons could arise from the annihilation of dark matter particles (such as a Kaluza–Klein particle⁷ with a mass of about 620 GeV).

High-energy electrons are rare. Before now, only emulsion chamber data have been available (above 10^{11} eV)^{4,8}, and these were of limited resolution, owing to the small depth of the calorimeters used, and low statistical significance, as the results were analysed manually⁹. Here we present new data from the Advanced Thin Ionization Calorimeter (ATIC) instrument^{10,11}, which contains a deep, fully active, bismuth germanate (BGO) calorimeter of 18 radiation lengths (X_0 , a characteristic distance for energy loss by bremsstrahlung) in eight layers arranged in orthogonal pairs to measure the energy deposited through a cascade (shower) of nuclear and electromagnetic interactions. At each step of the cascade, the energy of the primary particle is subdivided among many secondary particles. The area under the curve of ionization energy plotted against depth in the medium provides a measure of the incident particle energy, and the lateral distribution of energy across each layer can be used to separate electrons from protons. The topmost element is a pixelated silicon matrix (SiM) detector (4,480 pixels) which measures the charge of the incident particle. This is followed by three layers of scintillator hodoscopes embedded within a 30-cm-thick graphite target (this amount of material is 0.75 times the proton interaction length, and 1.5 times X_0). We determine the trajectory by using the hodoscopes along with cascade centroid positions in the BGO calorimeter.

The ATIC instrument has been studied extensively with GEANT and FLUKA simulations and was calibrated at the CERN SPS with proton and electron beams¹¹. The accelerator data validated the simulations, demonstrated an electron energy resolution of around 2% and verified the hadron–lepton separation capabilities^{11,12}. The detailed electron data analysis is described elsewhere^{12,13}, and is

reviewed briefly here and in the Supplementary Information (section 1). The basic ATIC energy calibration is provided by cosmic-ray muons recorded just before each flight, as well as by the shower data itself. The raw flight data are processed to physics units using these calibrations plus the temperature dependence of the BGO response.

The trajectory of each event is extrapolated to the SiM and identifies the pixels containing the primary particle signal. The SiM separates events with atomic number $Z \geq 2$ from the $Z = 0, 1$ events, and an absence of signal indicates a γ -ray candidate. The cascade profile is then analysed in the calorimeter. Electromagnetic cascades from electrons and γ -rays are narrower than hadronic cascades induced by interacting protons (whose products spread throughout the 30-cm target region). In addition, electrons and γ -rays deposit at least 85% of their energy in the calorimeter, whereas hadronic events deposit around 35%. Thus, at the bottom of the calorimeter the electron showers are dying out, whereas hadronic showers are usually still developing. As illustrated in Fig. 1, for electron candidates with energies over 50 GeV, in the histogram of Fig. 1, the electron peak (left of figure) is separated from the protons. Moreover, the atmospheric secondary γ -rays provide a calibration for the electron–proton separation (see Supplementary Information section 1 for details.)

For consistency checking, the data were analysed (1) in different zenith angle bins, (2) in different time periods corresponding to different trigger conditions, (3) in different physical sections of the apparatus and (4) with more severe geometrical cuts on the SiM and the calorimeter edge crystals. Furthermore, we carried out a ‘blind’ study in which all events were subject to the electron analysis. After correcting for the known differences in $Z \geq 2$ showers compared with protons, this analysis verified the electron–proton separation¹² (Supplementary Information section 2). A convolution of the rejection function with the measured proton energy deposit spectra provides the (energy-dependent) proton background. In Fig. 2 we show the raw spectrum of electron candidates at the instrument (multiplied by E^3 , where E is electron energy) for both the ATIC-1 and ATIC-2 flights, which are in excellent agreement. We also show the combined background (unresolved protons, misidentified γ -rays and atmospheric secondary electrons), which must be subtracted from the raw data.

After subtracting the background from each energy bin and correcting for energy loss in the overlying atmosphere, the absolute primary electron spectrum at the top of atmosphere is obtained (Fig. 3). Below 100 GeV, the ATIC spectrum agrees with previous data and with the ‘general’ spectrum calculated with the GALPROP interstellar propagation code¹⁴. Above about 100 GeV, the results

¹Purple Mountain Observatory, CAS, 2 West Beijing Road, Nanjing 210008, China. ²Max Planck Institute for Solar System Research, 2 Max Planck-Strasse, Katlenburg-Lindau 37191, Germany. ³Marshall Space Flight Center, Huntsville, Alabama 35812, USA. ⁴University of Maryland, Institute for Physical Science & Technology, College Park, Maryland 20742, USA. ⁵Skobel'syn Institute of Nuclear Physics, Moscow State University, Leninskie gory, GSP1, Moscow 119991, Russia. ⁶Louisiana State University, Department of Physics and Astronomy, Baton Rouge, Louisiana 70803, USA.

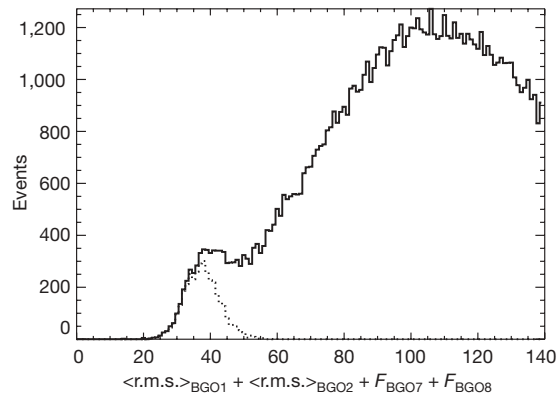


Figure 1 | Separation of electrons from protons in the ATIC instrument.

Candidate electron events (162,000) with energy over 50 GeV are plotted as a histogram with the horizontal axis showing the sum of the ‘weighted energy fraction’ (F values as defined below) in the last two BGO layers and the shower width (root mean squared, r.m.s.) in the first two layers. The shower width is calculated as

$$\langle r.m.s. \rangle^2 = \frac{\sum_{i=1}^n E_i (X_i - X_c)^2}{\sum_{i=1}^n E_i}$$

where X_c is the coordinate of the energy centre, X_i is the coordinate of the centre of the i th crystal and E_i is the energy deposited in the i th crystal. The F value is calculated as $F_n = (E_n / \text{Sum}) \langle r.m.s. \rangle^2$ where E_n is the energy deposit in BGO layer n , Sum is the total energy deposit in all BGO layers and $\langle r.m.s. \rangle$ refers to layer n (ref. 12). Each event is also fitted to an electromagnetic cascade profile to estimate the starting point and the depth of the cascade maximum. An event is accepted if the cascade starts above the first BGO layer, which eliminates many protons ($\sim 75\%$) but passes most electrons ($\sim 90\%$). Next a diagonal cut in r.m.s. and F is determined for each energy bin and used to isolate the electrons. This removes most of the protons (2 in 10^4 remain) and retains 84% of the electrons¹². The selected electrons are shown as the dotted histogram.

depart from the calculated curve. They show an excess electron flux up to about 650 GeV, above which the spectrum drops rapidly, with a return to the ‘general’ spectrum line at ~ 800 GeV. In particular, over the energy range 300 to 800 GeV we observe 210 electrons, whereas GALPROP predicts only 140 events, an excess of about six standard deviations. Using a source-on/source-off method for determining ‘significance’¹⁵, we obtain an excess of roughly four standard deviations (Supplementary Information section 4).

Data recently became available from the Polar Patrol Balloon (Antarctic) flight of the BETS detector. Although of lower statistical precision, results from the PPB-BETS calorimeter¹⁶ also indicate a possible structure and agree with the ATIC results (see Fig. 3), giving added confidence to the conclusion that this feature is real.

We varied the source injection parameters in the GALPROP code to try to reproduce the data points at 500 to 700 GeV. This required a hard injection spectrum which could not reproduce the drop in flux above 650 GeV and led to overproducing electrons above 1 TeV by a factor of almost three (and underproducing the well-measured data below 100 GeV).

The observed electron ‘feature’ therefore indicates a nearby source of high-energy electrons. This may be the result of an astrophysical object, as energetic electrons have been observed in a variety of astrophysical sites (for example in a supernova remnant¹⁷, pulsar wind nebula^{5,18}, micro-quasar⁶ or accreting intermediate-mass black hole). To fit the electron excess, such a source would need a very steep energy spectrum (spectral index around -1.4) with a high-energy cut-off at about 600–700 GeV, so as not to overproduce teraelectron-volt electrons. It is possible that a micro-quasar could produce a sharp feature in the electron spectrum⁶, but such an object would need to be local (less than 1 kpc away) and active relatively recently. Imaging atmospheric Cherenkov telescopes have observed numerous

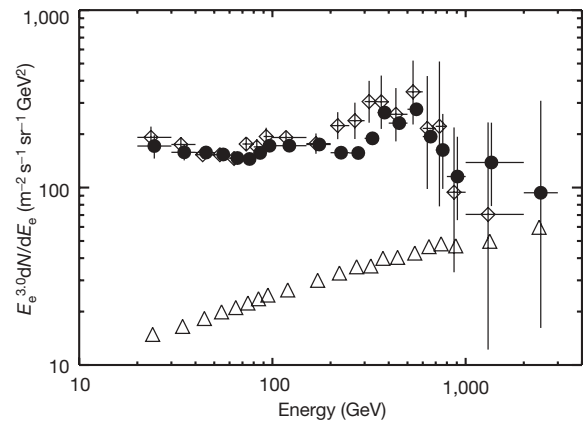


Figure 2 | ATIC-1 and ATIC-2 spectra at balloon altitude, showing good agreement with each other. The measured primary electron flux (scaled by E^3) at flight altitude is shown for ATIC-1 (open squares) and ATIC-2 (filled circles). The errors are one standard deviation. Both balloon flights were from McMurdo, Antarctica, and circumnavigated that continent. ATIC-1 was a test flight in 2000–01 and the usable data correspond to an exposure of $0.61 \text{ m}^2 \text{ sr days}$. ATIC-2 was a science flight in 2002–03 with an exposure of $2.47 \text{ m}^2 \text{ sr days}$. To eliminate edge effects, we restrict the incident zenith angle to be less than $\sim 37^\circ$ ($\cos \theta \geq 0.8$), use only the central 80% of the SiM and eliminate events in the outer crystals in each BGO layer. Within these limits, the electron detection efficiency above 60 GeV is 84% essentially independent of energy. The effective acceptance was determined as a function of particle energy considering the trigger efficiency, trajectory reconstruction efficiency and the geometrical restrictions. The effective acceptance of the instrument increases from $0.075 \text{ m}^2 \text{ sr}$ at 20 GeV to $0.15 \text{ m}^2 \text{ sr}$ for $E > 60$ GeV. Above 100 GeV, a total of 1,724 electron events were observed, with the highest energy event at 2.3 TeV. The total background is also shown in the figure as the open triangles and is a combination of unresolved protons, unidentified γ -rays and atmospheric secondary electrons produced in the material ($\sim 4.5 \text{ g cm}^{-2}$) above the instrument. ATIC becomes background limited for electrons only above several teraelectronvolts.

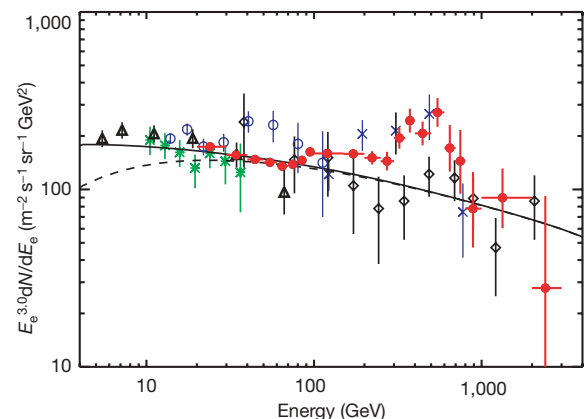


Figure 3 | ATIC results showing agreement with previous data at lower energy and with the imaging calorimeter PPB-BETS at higher energy. The electron differential energy spectrum measured by ATIC (scaled by E^3) at the top of the atmosphere (red filled circles) is compared with previous observations from the Alpha Magnetic Spectrometer AMS (green stars)³¹, HEAT (open black triangles)³⁰, BETS (open blue circles)³², PPB-BETS (blue crosses)¹⁶ and emulsion chambers (black open diamonds)^{4,8,9}, with uncertainties of one standard deviation. The GALPROP code calculates a power-law spectral index of -3.2 in the low-energy region (solid curve)¹⁴. (The dashed curve is the solar modulated electron spectrum and shows that modulation is unimportant above ~ 20 GeV.) From several hundred to ~ 800 GeV, ATIC observes an ‘enhancement’ in the electron intensity over the GALPROP curve. Above 800 GeV, the ATIC data returns to the solid line. The PPB-BETS data also seem to indicate an enhancement and, as discussed in Supplementary Information section 3, within the uncertainties the emulsion chamber results are not in conflict with the ATIC data.

sources of high-energy gamma radiation, including pulsar wind nebula and supernova remnants, finding spectra that fall as E^{-2} with an exponential cut-off in the teraelectronvolt region^{3,17}. This implies that particles have been accelerated to tens of teraelectronvolts or higher, which would not be consistent with the current electron data. The nearby object that comes closest to meeting the source requirements is the Geminga pulsar and associated remnant, whose potential contribution to high-energy electrons has been modelled^{4,18}. However, the calculated flux from Geminga is about a factor of 60 too low to explain the observations (see Supplementary Information section 5). Nevertheless, the classes of object discussed here have the potential to produce energetic electrons, and there may well be a nearby, unstudied astrophysical object that is accelerating the electrons observed by ATIC.

An alternative explanation invokes annihilation of dark matter particles. There has been considerable theoretical work on the predicted dark matter distribution in the Galaxy as well as on the production and propagation of the products of dark matter annihilations^{19–23}. Electrons and positrons are predicted as products of the annihilation of some exotic particles suggested as dark matter candidates²⁴, including weakly interacting particles from supersymmetric theories, such as neutralinos, and particles resulting from theories involving compactified extra dimensions—the ‘Kaluza–Klein’ (KK) particles⁷. The annihilation of supersymmetric and Kaluza–Klein types of dark matter can proceed through different channels including production of either electron–positron pairs or high-energy γ -rays (Supplementary Information section 6). The signature of this annihilation process is an increase in electron intensity above that expected from astrophysical sources, the details of which depend on the dark matter type and primary annihilation channel. Direct production of e^+e^- pairs is suppressed for supersymmetric particles, resulting in a source spectrum that has a broad peak and decreases in flux up to the particle mass¹⁹. This spectrum is further broadened by propagation and would not be consistent with the electron data. In contrast, direct production of e^+e^- pairs is not suppressed for Kaluza–Klein particles, resulting in a source spectrum that is dominated by a delta function at the particle mass. Energy losses during propagation broaden this distribution to lower energies. According to current theory, the mass of the lightest Kaluza–Klein particle is expected to be greater than 300 GeV (refs 19, 20). Further, the Wilkinson Microwave Anisotropy Probe (WMAP) has observed an excess in the microwave emission around the inner region of our Galaxy (‘WMAP haze’) that could be a product of dark matter annihilation. This assumption provides a constraint on the dark matter annihilation rate^{19,23}. For Kaluza–Klein particles, the annihilation rate is inversely proportional to the square of the particle mass, and the mass of the lightest Kaluza–Klein particle that could reproduce the WMAP haze is estimated to be 550 to 650 GeV (refs 25, 26).

The GALPROP code includes the capability to inject and propagate a source of electrons resulting from the annihilation of a dark matter particle^{14,21}. As an example, the spectrum produced for a Kaluza–Klein particle mass of 620 GeV is shown in Fig. 4. When added to the general spectrum, this reproduces the observed data well. The ATIC energy range includes this mass and, therefore, the calculation should be relatively immune to uncertainties in the overall dark matter distribution, and to galactic propagation, but would be sensitive to conditions in our local galactic neighbourhood²². The difficulty is that a model with a smooth distribution of Kaluza–Klein particles annihilating in our Galaxy produces a much smaller signal than the feature reported here. To be consistent with the WMAP haze, the annihilation rate for a 620-GeV thermal relic Kaluza–Klein particle would need to be about $4.4 \times 10^{-26} \text{ cm}^3 \text{ s}^{-1}$, a factor of ~ 200 smaller than that required to fit the observed electron excess. Such enhancements are usually attributed to a ‘boost factor’ associated with non-uniform clumps in the dark matter distribution²⁷, and such clumps could also be located near our Solar System²⁸. Moreover, ‘minispikes’ of dark matter overdensities, associated for instance with

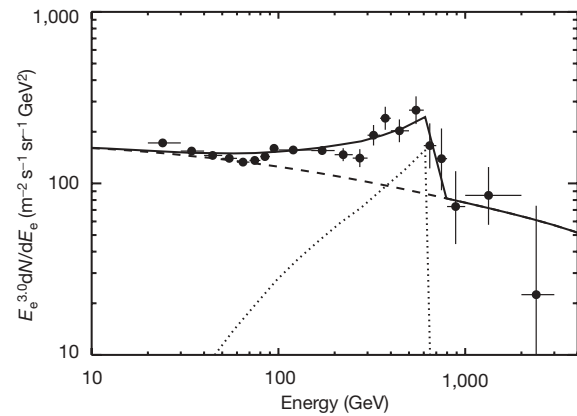


Figure 4 | Assuming an annihilation signature of Kaluza–Klein dark matter, all the data can be reproduced. The GALPROP general electron spectrum resulting from sources across the galaxy is shown as the dashed line. The dotted curve represents the propagated electrons from the annihilation of a Kaluza–Klein particle. The dotted curve assumes an isothermal dark matter halo of 4-kpc scale height, a local dark matter density of 0.43 GeV cm^{-3} , a Kaluza–Klein mass of 620 GeV, and an annihilation cross section rate of $1 \times 10^{-23} \text{ cm}^3 \text{ s}^{-1}$, which implies a boost factor of ~ 200 . The sum of these signals is the solid curve. Here the spectrum is multiplied by $E^{3.0}$ for clarity. The solid curve provides a good fit to both the magnetic spectrometer data^{30,31} and calorimeter data^{16,32} and reproduces all of the measurements from 20 GeV to 2 TeV, including the cut-off in the observed excess. All error bars are one standard deviation.

intermediate-mass black holes, can result in boost factors of a few thousand²⁹. In any case, the exact level of ‘boost’ is still subject to debate.

It should be noted that other authors^{19,21} have found the need to introduce boost factors of 200–300 to explain the cosmic-ray positron excess observed by the HEAT magnetic spectrometer experiment³⁰ in terms of an annihilation signature of Kaluza–Klein dark matter. Thus, a model for Kaluza–Klein dark matter annihilation that would explain the observed ATIC electron excess could also fit the excess positrons observed by HEAT at $\sim 30 \text{ GeV}$.

The ‘feature’ in the spectrum of high-energy cosmic-ray electrons reported here provides an intriguing puzzle. Either an as yet unknown astrophysical source or the annihilation of a dark matter particle is a possible explanation. If the ‘feature’ is caused by an astrophysical object this would be the first direct observation of a nearby source of particles with energies of hundreds of gigaelectronvolts and would open a new window for studying such objects. Kaluza–Klein dark matter arises from multi-dimensional theories of our Universe in which the extra dimensions are ‘compact’, meaning that they have only a small (but non-zero) effect on our four-dimensional physical reality. If the Kaluza–Klein annihilation explanation proves to be correct, this will necessitate a fuller investigation of such multi-dimensional spaces, with potentially important implications for our understanding of the Universe.

Received 23 May; accepted 1 October 2008.

1. Tang, K.-K. The energy spectrum of electrons and cosmic-ray confinement. A new measurement and its interpretation. *Astrophys. J.* **278**, 881–892 (1984).
2. Allen, G. E. *et al.* Evidence of X-ray synchrotron emission from electrons accelerated to 40 TeV in the supernova remnant Cassiopeia A. *Astrophys. J.* **487**, L97–L100 (1997).
3. Aharonian, F. A. *et al.* High-energy particle acceleration in the shell of a supernova remnant. *Nature* **432**, 75–77 (2004).
4. Kobayashi, T. *et al.* The most likely sources of high-energy cosmic-ray electrons in supernova remnants. *Astrophys. J.* **601**, 340–351 (2004).
5. Aharonian, F. A. *et al.* Cosmic ray positrons connected with galactic gamma radiation of high and very high energies. *J. Phys. G Nucl. Part. Phys.* **17**, 1769–1778 (1991).
6. Heinz, S. & Sunyaev, R. Cosmic rays from microquasars: A narrow component in the CR spectrum. *Astron. Astrophys.* **390**, 751–766 (2002).

7. Cheng, H. C., Feng, J. L. & Matchev, K. T. Kaluza–Klein dark matter. *Phys. Rev. Lett.* **89**, 211301 (2002).
8. Nishimura, J. *et al.* in *25th ICR Conf. Papers* Vol. 4 (eds Potgieter, M. S., Raubenheimer, B. C. & van der Walt, O. J.) 233–237 (Space Research Unit, Durban, 1997).
9. Nishimura, J. *et al.* Emulsion chamber observations of primary cosmic-ray electrons in the energy range 30–1000 GeV. *Astrophys. J.* **238**, 394–409 (1980).
10. Guzik, T. G. *et al.* The ATIC long duration balloon project. *Adv. Space Res.* **33**, 1763–1770 (2004).
11. Ganel, O. *et al.* Beam tests of the balloon-borne ATIC experiment. *Nucl. Instrum. Methods A* **552**, 409–419 (2005).
12. Chang, J. *et al.* Resolving electrons from protons in ATIC. *Adv. Space Res.* **42**, 431–436 (2008).
13. Chang, J. *et al.* in *Proc. 28th Intl Cosmic Ray Conf.* (eds Kajita, T., Asaoka, Y., Kawachi, A., Matsubara, Y. & Sasaki, M.) 1817–1820 (Universal Acad. Press, Tokyo, 2003).
14. Strong, A. W. & Moskalenko, I. V. Models for galactic cosmic-ray propagation. *Adv. Space Res.* **27**, 717–726 (2001).
15. Li, T. P. & Ma, Y.-Q. Analysis methods for results in gamma-ray astronomy. *Astrophys. J.* **272**, 317–324 (1983).
16. Torii, S. *et al.* High energy electron observation by Polar Patrol Balloon flight in Antarctica. *Adv. Polar Upper Atmos. Res.* **20**, 52–62 (2006).
17. Aharonian, F. A. *et al.* Very high energy gamma rays from the composite SNR G 0.9+0.1. *Astron. Astrophys.* **432**, L25–L29 (2005).
18. Büsching, I., de Jager, O. C., Potgieter, M. S. & Venter, C. A cosmic-ray positron anisotropy due to two middle-aged, nearby pulsars? *Astrophys. J.* **678**, L39–L42 (2008).
19. Hooper, D. & Silk, J. Searching for dark matter with future cosmic positron experiments. *Phys. Rev. D* **71**, 083503 (2005).
20. Applequist, T. & Yee, H.-U. Universal extra dimensions and the Higgs boson mass. *Phys. Rev. D* **67**, 055002 (2003).
21. Moskalenko, I. V. & Strong, A. W. Galactic propagation of positrons from particle dark-matter annihilation. *Phys. Rev. D* **60**, 063003 (1999).
22. Delahaye, T. *et al.* Positrons from dark matter annihilation in the galactic halo: Theoretical uncertainties. *Phys. Rev. D* **77**, 063527 (2008).
23. Hooper, D. Indirect searches for dark matter. Preprint at (<http://arXiv.org/abs/0710.2062v1>) (2008).
24. Drees, M. & Gerbier, G. Supersymmetric dark matter. *Phys. Lett. B* **592**, 216–219 (2004).
25. Servant, G. & Tait, T. M. P. Is the lightest Kaluza–Klein particle a viable dark matter candidate? *Nucl. Phys. B* **650**, 391–419 (2003).
26. Hooper, D., Finkbeiner, D. P. & Dobler, G. Possible evidence for dark matter annihilation from excess microwave emission around the center of the Galaxy seen by the Wilkinson Microwave Anisotropy Probe. *Phys. Rev. D* **76**, 083012 (2007).
27. Lavalle, J. *et al.* Full calculation of clumpiness boost factors for antimatter cosmic rays in the light of Λ CDM N -body simulation results: Abandoning hope in clumpiness enhancement? *Astron. Astrophys.* **479**, 427–452 (2008).
28. Diemand, J. *et al.* Clumps and streams in the local dark matter distribution. *Nature* **454**, 735–738 (2008).
29. Brun, P. *et al.* Antiproton and positron signal enhancement in dark matter minispikes scenarios. *Phys. Rev. D* **76**, 083506 (2007).
30. Barwick, S. W. *et al.* Measurements of the cosmic-ray positron fraction from 1 to 50 GeV. *Astrophys. J.* **482**, L191–L194 (1997).
31. Alcaraz, J. *et al.* Leptons in near Earth orbit. *Phys. Lett. B* **484**, 10–22 (2000).
32. Torii, S. *et al.* The energy spectrum of cosmic-ray electrons from 10 to 100 GeV observed with a highly granulated imaging calorimeter. *Astrophys. J.* **559**, 973–984 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This research was supported in the USA by NASA, in Russia by the Russian Foundation for Basic Research, and in China by the National Natural Science Foundation. The help of the NASA BPO and CSBF during balloon flights and the US NSF and RPSC for Antarctica operations is gratefully acknowledged.

Author Contributions Each institution in the ATIC collaboration assumed particular roles, each of which provided important contributions to this paper. In particular, the Max Planck Institute and Purple Mountain Observatory (PMO) developed the electron observation techniques. Marshall Space Flight Center (MSFC), Moscow State University (MSU) and Louisiana State University (LSU) developed the hardware which was integrated at LSU, who also led the flight team. The University of Maryland (UMD) and MSU performed instrument simulations and processed the data. PMO led the electron analysis of the flight data and LSU and PMO prepared the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.P.W. (wefel@phunds.phys.lsu.edu).

LETTERS

Isotropic quantum scattering and unconventional superconductivity

T. Park^{1,2}, V. A. Sidorov^{1,3}, F. Ronning¹, J.-X. Zhu¹, Y. Tokiwa¹, H. Lee¹, E. D. Bauer¹, R. Movshovich¹, J. L. Sarrao¹ & J. D. Thompson¹

Superconductivity without phonons has been proposed for strongly correlated electron materials that are tuned close to a zero-temperature magnetic instability of itinerant charge carriers¹. Near this boundary, quantum fluctuations of magnetic degrees of freedom assume the role of phonons in conventional superconductors, creating an attractive interaction that ‘glues’ electrons into superconducting pairs. Here we show that superconductivity can arise from a very different spectrum of fluctuations associated with a local (or Kondo-breakdown) quantum critical point^{2–5} that is revealed in isotropic scattering of charge carriers and a sublinear, temperature-dependent electrical resistivity. At this critical point, accessed by applying pressure to the strongly correlated, local-moment antiferromagnet CeRhIn₅, magnetic and charge fluctuations coexist and produce electronic scattering that is maximal at the optimal pressure for superconductivity. This previously unanticipated source of pairing glue⁶ opens possibilities for understanding and discovering new unconventional forms of superconductivity.

In conventional superconductors, excitations of a solid’s crystal lattice provide an attractive interaction that binds itinerant electrons into pairs with zero net spin and momentum⁷. The resulting state of superconductivity is very sensitive to the presence of paramagnetic impurities, such as cerium: coupling of the paramagnetic spin **S** to the spin **s** of itinerant electrons, in the form $J\mathbf{S}\cdot\mathbf{s}$, where J is the coupling strength, scatters electrons and breaks electron pairs⁸. Consequently, it has been a challenge to explain discoveries of unconventional superconductivity in strongly correlated electron materials with dense arrays of paramagnetic ions, such as cerium-based heavy-fermion compounds⁹. This form of spin–spin coupling is the origin of the Kondo effect, which gives itinerant charge carriers their heavy mass in heavy-fermion systems and, in the process, creates a ‘large’ Fermi volume that encloses the paramagnetic electrons, for example the 4*f* electrons of cerium¹⁰.

The proximity of superconductivity in heavy fermions, such as the prototype CeCu₂Si₂, to a spin instability of their large Fermi surfaces suggests that magnetic fluctuations provide the attractive ‘glue’ in the same way that lattice fluctuations do in conventional superconductors^{1,9}. As a spin-density instability is tuned by a non-thermal parameter, such as pressure, towards a continuous transition at $T = 0$ K, magnetic excitations become quantum mechanically critical, leading to a divergent magnetic susceptibility at the ordering wave vector **Q**, a condition favourable for creating superconducting electron pairs with finite angular momenta⁹. Quantum fluctuations of the spin density also strongly scatter electrons on parts of the Fermi surface connected by the wave vector **Q**, giving an electrical resistivity, ρ , that increases from $T = 0$ K with a power-law dependence $\rho \propto T^\varepsilon$, where $1 \leq \varepsilon < 3/2$, which is distinct from the characteristic T^2 dependence

of a Fermi liquid that emerges as the strongly correlated electron system is tuned away from its quantum critical point¹¹.

Quantum criticality of a very different kind has been proposed for some heavy-fermion materials whose magnetism derives from the indirect interaction among localized magnetic moments^{2–5}. Deep in the magnetically ordered phase, the Kondo effect is absent and the Fermi volume is small, that is, does not include the localized electrons, but as the $T = 0$ boundary between paramagnetic and localized magnetic states is approached, the Fermi volume jumps in size. A consequence of such a quantum phase transition is that the Fermi surface becomes critical at the $T = 0$ magnetic–paramagnetic boundary, which is consistent with the criticality having a spatially local character and creates a spectrum of fluctuations that dominate physical properties, such as resistivity, even at temperatures well above $T = 0$ K (ref. 12). In comparison with a spin-density quantum phase transition that affects only a portion of the Fermi surface and involves only magnetic degrees of freedom, a local type of quantum phase transition is more ‘violent’, producing fluctuations in spin and charge channels that separately or collectively could promote an attractive interaction between electron pairs. However, the lack of a theoretical basis for superconductivity and the absence of superconductivity in the prime candidates for local criticality, YbRh₂Si₂ and CeCu_{6–x}Au_x, question the viability of superconductivity mediated by locally critical fluctuations⁶. Using the heavy-fermion antiferromagnet CeRhIn₅ as an example, we report evidence that fluctuations from a local form of quantum criticality can provide a new route to unconventional superconductivity.

Figure 1 displays a temperature–pressure map of the local exponent, $\varepsilon \equiv \partial \ln(\rho(T) - \rho(T = 0)) / \partial (\ln T)$, of the electrical resistivity of CeRhIn₅. Funnel-shaped yellow boundaries in Fig. 1 that emerge near the maximum pressure-induced superconducting transition temperature, T_c , mark a lower crossover temperature from an unusual state with a resistivity, $\rho \propto T^{0.85}$, sublinearly dependent on T (or ‘sub- T -linear’). (Representative pressure- and temperature-dependent resistivity curves and their analysis, shown in Supplementary Figs 1 and 2, document the origin of the colour map and phase boundaries.) In the low-pressure limit, the yellow boundary reflects the onset of short-ranged antiferromagnetic spin correlations over a small interval above the Néel temperature, T_N (ref. 13), but signals a change in resistivity to a $T^{3/2}$ dependence before settling into a heavy Fermi liquid state at $T < T_{FL}$ in the high-pressure limit. The upper boundary of this unusual state is determined by $T_{max}/2$, where T_{max} is the temperature at which the resistivity is a maximum. At 2.35 GPa, where there is no magnetic order, this phase with sub- T -linear resistivity extends to progressively lower temperatures as superconductivity is suppressed by an applied magnetic field. At 10 T, which is slightly higher than $H_{c2}(0)$, the field necessary to suppress superconductivity

¹Los Alamos National Laboratory, Los Alamos, New Mexico 87545, USA. ²Department of Physics, Sungkyunkwan University, Suwon 440-746, South Korea. ³Verezhchagin Institute of High Pressure Physics, RAS, 142190 Troitsk, Russia.

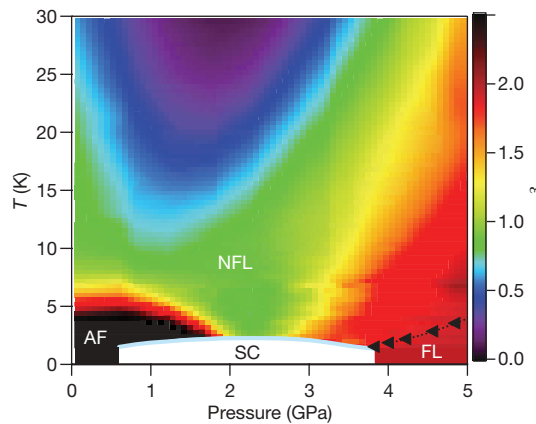


Figure 1 | Temperature–pressure phase diagram of CeRhIn₅. Colours represent the local exponent, $\varepsilon \equiv \partial(\ln \Delta\rho)/\partial(\ln T)$, at zero magnetic field, where $\Delta\rho = \rho - \rho(T = 0 \text{ K}) = AT^\varepsilon$. The resistivity, ρ , was measured along the c axis of this tetragonal crystal. The residual resistivity, $\rho(T = 0 \text{ K})$, and temperature coefficient, A , were obtained with least-squares fitting to the power-law form at low temperatures and are shown in Supplementary Fig. 1. Also shown are boundaries of the phase of local-moment antiferromagnetic (AF) order, the superconductivity (SC) phase, and the region of temperature below which the resistivity follows a T^2 temperature dependence characteristic of a Landau Fermi liquid. The cone-shaped region of green denotes a state of sub- T -linear resistivity, labelled NFL and discussed in the text, that appears to emanate from the dome of superconductivity, where $T_c = 2.3 \text{ K}$ is a maximum. For comparison, the SC and AF phases are coloured white and black, respectively.

completely, the weakly field-dependent sub- T -linearity holds from $\sim 10 \text{ K}$ to 250 mK , and crosses over to a T^2 (Fermi liquid) dependence below 150 mK (Supplementary Fig. 3). We emphasize that this sub- T -linear behaviour with a single temperature exponent over a broad temperature range is distinct from crossover behaviour that is often observed in heavy-fermion compounds as they are warmed from a low-temperature Fermi liquid behaviour (T^2 dependence) to a resistivity maximum near 100 K : the temperature exponent in the crossover regime varies with temperature; that is, there is no unique power-law behaviour.

Studies of specific heat as a function of pressure have revealed that the Néel transition extends continuously to $T = 0$ at 2.35 GPa in the presence of a magnetic field suppressing superconductivity¹⁴. This magnetic quantum critical point at 2.35 GPa offers a possible explanation for the strange metallic behaviour found at the point of maximum T_c ; however, the sub- T -linear resistivity is not anticipated by any theory of spin-density criticality¹⁵. The nonconformity of CeRhIn₅ to the spin-density model indicates that the quantum critical point in CeRhIn₅ may differ in nature from a magnetic instability of the Fermi surface. When coupled with quantum oscillation measurements at 2.35 GPa (ref. 16), which show an abrupt change in the Fermi surface, the convergence of multiple energy scales represented by the crossover and phase-transition boundaries in Fig. 1 is consistent with an interpretation as a local form of criticality that involves magnetic as well as fermionic degrees of freedom^{2–6}, where, now, the yellow boundary at high pressures can be understood as a crossover temperature below which incoherent scattering due to quantum fluctuations starts to yield to coherent scattering within the lattice of Kondo sites.

Strong support for the spatially local nature of this criticality comes from electrical transport measured perpendicular (ρ_{ab}) and parallel to the tetragonal c axis (ρ_c) (Fig. 2). At temperatures higher than $\sim 100 \text{ K}$, where charge scattering comes predominantly from randomly oriented cerium $4f$ -electron moments, the ratio $\rho_{ab}/\rho_c \approx 0.5$ is essentially independent of pressure and set primarily by intrinsic crystalline anisotropy. On the other hand, below T_{FL} and at high pressures, transport anisotropy increases by a factor of 2.5, to

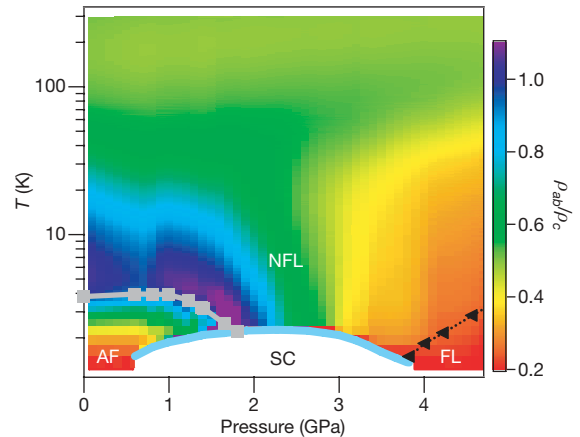


Figure 2 | Temperature–pressure variation of resistivity anisotropy.

Apparent structure in the colour map (describing ρ_{ab}/ρ_c) is due to an interpolation of data obtained at pressures slightly different from those used to determine ρ_{ab} and ρ_c . Significantly, the resistivity anisotropy found at high temperatures persists to the lowest temperature in the NFL pressure range where ρ_c (Fig. 1) and ρ_{ab} exhibit a sub- T -linear variation. Data representative of those from which this map was constructed are shown in Supplementary Fig. 4. The SC phase is coloured white.

give $\rho_{ab}/\rho_c \approx 0.2$. We ascribe this increase in anisotropy to characteristics of the large Fermi volume that now includes the ‘Kondo-ized’ $4f$ electrons of cerium. The related superconductor CeCoIn₅, which has a large Fermi volume, has the effective mass anisotropy $m_a^*/m_c^* \approx 0.18$ (ref. 17), where m^* reflects many-body correlation effects that also are present in CeRhIn₅ below T_{FL} .

The pronounced anisotropy observed in the Fermi liquid state becomes progressively less obvious with decreasing pressure. In the quantum critical regime, where the resistivity is sub- T -linear, the anisotropy ratio is almost temperature independent and is similar to that at room temperature. With the assumption of independent scattering sources, the total scattering rate is $1/\tau = 1/\tau_i + 1/\tau_q$, where τ_i and τ_q are the collision times for scattering by impurities and by excitations, respectively. Potential scattering from impurities is isotropic, leaving crystal anisotropy, which changes very weakly with pressure¹⁸, and scattering from fluctuations associated with the quantum phase transition as sources of resistive anisotropy. The absence of new anisotropy in the quantum critical regime argues against a spin-density quantum critical interpretation because, in this case, the strong scattering from parts of the Fermi surface spanned by \mathbf{Q} is expected to be highly anisotropic¹⁹, reflecting the Fermi surface topology of the large-volume paramagnetic state. Instead, isotropic scattering induced over the entire Fermi surface at a local or Kondo-breakdown quantum critical point is consistent with our observations. Model calculations, presented in the Supplementary Fig. 5, show that the sublinear temperature dependence of electrical resistivity can arise from isotropic scattering from fluctuations associated with a local quantum critical point. We note that a similar sub- T -linear resistivity has been shown to be present in a Kondo-breakdown model⁵.

Fluctuations from a local form of quantum criticality may be responsible for the unconventional superconductivity of CeRhIn₅. Figure 3 displays a map of the isothermal resistivity, normalized by the resistivity in the normal metallic phase at 5.2 GPa , as a function of pressure. As seen in this figure and in Supplementary Fig. 6, the highest scattering rate and the highest T_c occur simultaneously, which appears to contradict the conventional view that scattering is harmful to unconventional superconductivity^{20,21}: superconductivity disappears when the resistivity reaches a comparable value of $20 \mu\Omega \text{ cm}$ in disordered CeCoIn₅. Unlike chemical substitution, pressure does not induce extra disorder. Instead, the maximum resistivity in the vicinity of the quantum critical point is due to the build-up of quantum fluctuations that amplifies scattering by the small number of impurities in this very

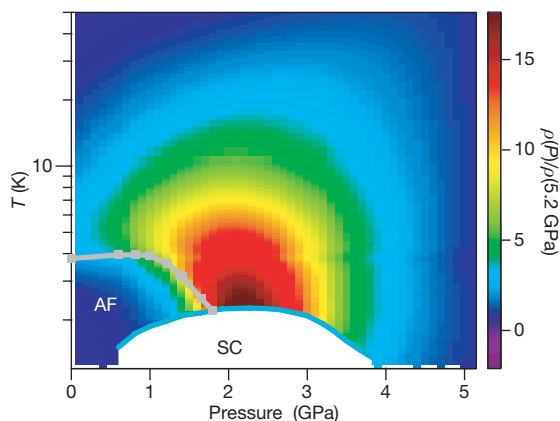


Figure 3 | Pressure-dependent c -axis resistivity. The colour map describes the c -axis resistivity normalized by its value in the normal metallic state at 5.2 GPa. This resistivity ratio is highest at temperatures and pressures where the resistivity is sub- T -linear (Fig. 1) and resistivity anisotropy (Fig. 2) is the same as at high temperatures. Maximum scattering appears above the maximum T_c of the superconducting dome. Representative isothermal cuts of the resistivity ratio, used to construct the colour map, are shown in Supplementary Fig. 6. The SC phase is coloured white.

pure single crystal of CeRhIn₅. The high spectral density of quantum fluctuations, reflected in the maximum of ρ_c (which occurs at 2.3 K), then provides the glue for optimal electron pairing.

In CeCu₂Si₂, where two domes of superconductivity emerge with applied pressure, spin-density and valence fluctuations are proposed to mediate electron pairing separately for each dome, with a T_c higher by a factor of four in the dome associated with valence fluctuations²². In CeRhIn₅, the local nature of quantum criticality implies simultaneous fluctuations in spin and charge channels. Our experiments cannot resolve which channels or channel dominate(s) the pairing interaction in CeRhIn₅, but the multi-criticality of CeRhIn₅ appears to be key to understanding the sub- T -linear resistivity and the robustness of the superconductivity over a wide pressure range. This work raises the fundamental issues of determining (1) how multiple fluctuating channels interplay with each other to produce superconductivity; (2) the appropriate description of fluctuations that arise from a local or Kondo-breakdown quantum critical point; (3) why there is no superconductivity in comparably pure crystals of YbRh₂Si₂; and (4) whether a similar analysis could be applicable to other heavy-fermion superconductors such as PuCoGa₅, where composite pairs of local moments and electrons are proposed²³ to condense to form superconductivity. The solutions of these problems will guide the search for new examples of unconventional superconductivity and will be broadly applicable to other strongly correlated superconductors, such as the high- T_c copper oxide superconductors²⁴, in which two domes of superconductivity or one extensive superconducting dome may naturally arise, as a function of doping, from multiple quantum critical points.

Received 16 June; accepted 10 September 2008.

1. Monthoux, P., Pines, D. & Lonzarich, G. G. Superconductivity without phonons. *Nature* **450**, 1177–1183 (2007).

2. Si, Q., Rabello, S., Ingersent, K. & Smith, J. L. Locally critical quantum phase transitions in strongly correlated metals. *Nature* **413**, 804–808 (2001).
3. Coleman, P., Pepin, C., Si, Q. & Ramazashvili, R. How do Fermi liquids get heavy and die? *J. Phys. Condens. Matter* **13**, R723–R738 (2001).
4. Paul, I., Pepin, C. & Norman, M. R. Kondo breakdown and hybridization fluctuations in the Kondo-Heisenberg lattice. *Phys. Rev. Lett.* **98**, 026402 (2007).
5. Pepin, C. Kondo breakdown as a selective Mott transition in the Anderson lattice. *Phys. Rev. Lett.* **98**, 206401 (2007).
6. Gegenwart, P., Si, Q. & Steglich, F. Quantum criticality in heavy-fermion metals. *Nature Phys.* **4**, 186–197 (2008).
7. Bardeen, J., Cooper, L. N. & Schrieffer, J. R. Theory of superconductivity. *Phys. Rev.* **108**, 1175–1204 (1957).
8. Abrikosov, A. A. & Gor'kov, L. P. Contribution to the theory of superconducting alloys with paramagnetic impurities. *Zh. Eksp. Teor. Fiz.* **39**, 1781–1796 (1960); *Sov. Phys. JETP* **12**, 1243–1253 (1961).
9. Mathur, N. D. *et al.* Magnetically mediated superconductivity in heavy fermion compounds. *Nature* **394**, 39–43 (1998).
10. Hewson, A. C. *The Kondo Problem to Heavy Fermions* (Cambridge Univ. Press, 1993).
11. Rosch, A. Interplay of disorder and spin fluctuations in the resistivity near a quantum critical point. *Phys. Rev. Lett.* **82**, 4280–4283 (1999).
12. Paschen, S. *et al.* Hall-effect evolution across a heavy-fermion quantum critical point. *Nature* **432**, 881–885 (2004).
13. Bao, W. *et al.* Anisotropic three-dimensional magnetic fluctuations in heavy fermion CeRhIn₅. *Phys. Rev. B* **65**, 100505 (2002).
14. Park, T. *et al.* Hidden magnetism and quantum criticality in the heavy fermion superconductor CeRhIn₅. *Nature* **440**, 65–68 (2006).
15. Stewart, G. R. Non-Fermi-liquid behavior in d- and f-electron metals. *Rev. Mod. Phys.* **73**, 797–885 (2001).
16. Shishido, H. *et al.* A drastic change of the Fermi surface at a critical pressure in CeRhIn₅: dHvA study under pressure. *J. Phys. Soc. Jpn* **74**, 1103–1106 (2005).
17. Settai, R. *et al.* De Haas-van Alphen experiments under extreme conditions of low temperature, high field and high pressure, for high-quality cerium and uranium compounds. *J. Phys. Condens. Matter* **13**, L627–L634 (2001).
18. Normile, P. S. *et al.* High-pressure structural parameters of the superconductors CeMnIn₅ and PuMnGa₅ (M=Co, Rh). *Phys. Rev. B* **72**, 184508 (2005).
19. Lohneysen, H. V., Rosch, A., Vojta, M. & Wolfe, P. Fermi-liquid instabilities at magnetic quantum phase transitions. *Rev. Mod. Phys.* **79**, 1015–1075 (2007).
20. Mackenzie, A. P. *et al.* Extremely strong dependence of superconductivity on disorder in Sr₂RuO₄. *Phys. Rev. Lett.* **80**, 161–164 (1998).
21. Paglione, J., Sayles, T. A., Ho, P.-C., Jeffries, J. R. & Maple, M. B. Incoherent non-Fermi-liquid scattering in a Kondo lattice. *Nature Phys.* **3**, 703–706 (2007).
22. Yuan, H. Q. *et al.* Observation of two distinct superconducting phases in CeCu₂Si₂. *Science* **302**, 2104–2107 (2003).
23. Flint, R., Dzero, M. & Coleman, P. Heavy electrons and the symplectic symmetry of spin. *Nature Phys.* **4**, 643–648 (2008).
24. Panagopoulos, C. *et al.* Evidence for a generic quantum transition in high- T_c cuprates. *Phys. Rev. B* **66**, 064501 (2002).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements The authors thank Q. Si, C. D. Batista, A. V. Balatsky, C. Varma, Z. Nussinov, D. Pines and N. J. Curro for discussions. Work at Los Alamos National Laboratory was performed under the auspices of the US Department of Energy, Office of Science, with support from the Los Alamos Directed Research and Developmental programme. V.A.S. appreciates the support of the Russian Foundation for Basic Research (grant no. 06-02-16590) and the Program of the Presidium of RAS on Physics of Strongly Compressed Matter.

Author Contributions T.P., V.A.S., F.R., Y.T., H.L. and R.M. collected data. E.D.B. and J.L.S. synthesized CeRhIn₅ and LaRhIn₅ single crystals. J.-X.Z. and F.R. analysed data. T.P. and J.D.T. designed the study, analysed data and wrote the paper. All authors discussed the results and commented on the manuscript.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to T.P. (tuson@lanl.gov) or J.D.T. (jdt@lanl.gov).

Twinning superlattices in indium phosphide nanowires

Rienk E. Algra^{1,2,3}, Marcel A. Verheijen², Magnus T. Borgström^{2†}, Lou-Fé Feiner², George Immink², Willem J. P. van Enckevort³, Elias Vlieg³ & Erik P. A. M. Bakkers²

Semiconducting nanowires offer the possibility of nearly unlimited complex bottom-up design^{1,2}, which allows for new device concepts^{3,4}. However, essential parameters that determine the electronic quality of the wires, and which have not been controlled yet for the III–V compound semiconductors, are the wire crystal structure and the stacking fault density⁵. In addition, a significant feature would be to have a constant spacing between rotational twins in the wires such that a twinning superlattice is formed, as this is predicted to induce a direct bandgap in normally indirect bandgap semiconductors^{6,7}, such as silicon and gallium phosphide. Optically active versions of these technologically relevant semiconductors could have a significant impact on the electronics⁸ and optics⁹ industry. Here we show first that we can control the crystal structure of indium phosphide (InP) nanowires by using impurity dopants. We have found that zinc decreases the activation barrier for two-dimensional nucleation growth of zinc-blende InP and therefore promotes crystallization of the InP nanowires in the zinc-blende, instead of the commonly found wurtzite, crystal structure¹⁰. More importantly, we then demonstrate that we can, once we have enforced the zinc-blende crystal structure, induce twinning superlattices with long-range order in InP nanowires. We can tune the spacing of the superlattices by changing the wire diameter and the zinc concentration, and we present a model based on the distortion of the catalyst droplet in response to the evolution of the cross-sectional shape of the nanowires to quantitatively explain the formation of the periodic twinning.

Twin planes and, more generally, planar stacking faults are commonly found in III–V nanowires grown in the [111] direction by the vapour-liquid-solid (VLS) mechanism. A twin plane in a zinc-blende (stacking fault in a wurtzite) nanowire can be considered as a monolayer of the wurtzite (zinc blende) phase¹¹. Stacking faults can significantly affect the electronic properties of the nanowires^{5,7}. The electron wavefunction is discontinuous at a stacking fault, which leads, for instance, to a reduced mobility of charge carriers. The formation and resulting morphology of randomly distributed stacking faults in nanowires have been investigated by several authors^{11–14}. Twin planes that have a constant spacing within a nanowire form a twinning superlattice (TSL); this modifies the electronic band structure, giving rise to the formation of minibands⁷. Recently, small domain TSLs have been observed locally in bulk Si^{15,16}, and occasionally in ZnS nanowires¹⁷, but the parameters controlling the phenomenon were not identified.

The first step in obtaining a TSL is to control the nanowire crystal structure. Bulk InP has the zinc-blende crystal structure, because the free energy is slightly lower ($\Delta E = 6.8$ meV per III–V atom pair)¹⁸ for this structure than for InP with the wurtzite structure. However,

nominally undoped InP nanowires commonly exhibit the wurtzite crystal structure. Possible explanations for the formation of wurtzite nanowires are the lower surface energy of the parallel side facets of wurtzite wires compared to that of zinc-blende wires¹⁸ and the interface energies at the vapour-liquid-solid three-phase line¹⁹. These effects would make crystallization in the wurtzite phase especially favourable for thin wires that have a large surface to bulk ratio.

We have synthesized InP nanowires from colloidal gold particles by VLS growth using metal-organic vapour phase epitaxy (MOVPE) with trimethylindium and phosphine as molecular precursors (Methods Summary). We consistently observe that undoped InP wires (with diameters of 10–250 nm) have the wurtzite structure, though in general contain many stacking faults. With increasing diameter the number of stacking faults decreases, leading to wires with a larger fraction of wurtzite structure and showing that the above-mentioned factors do not ultimately determine the crystal structure of the wires.

The main difference between bulk and VLS growth is the presence of the catalyst particle from which the crystal is precipitated, and therefore the atomic interactions at the liquid-solid interface should be considered. We find that the parameter critically determining the nanowire crystal structure and the stacking fault density is the chemical composition of the catalyst particle near the liquid-solid interface. We introduce diethylzinc in the growth system to establish p-type doping in our nanowires. Importantly, when sufficient Zn is added to the system, the nanowires precipitate in the zinc-blende crystal structure. We find a transition from wurtzite to zinc-blende structure at a diethylzinc partial pressure of 4.6×10^{-5} mbar (4.6×10^{-4} mbar corresponds to a free hole concentration of 10^{18} cm⁻³ in the InP nanowires²⁰), as is shown in Fig. 1a (see also Supplementary Information 1). For diethylzinc partial pressures between 4.6×10^{-5} and 4.6×10^{-4} mbar, we find that twin planes are randomly distributed in the nanowires. Strikingly, above 4.6×10^{-4} mbar, the twin planes exhibit a constant spacing for a given Zn concentration and wire diameter, and the nanowire develops a TSL. The segment length of the TSL increases with the Zn concentration.

To quantify the effect of Zn on the crystal structure, we have calculated, based on two-dimensional nucleation, a kinetic phase diagram (Fig. 1b). The obtained curve, which relates the supersaturation in the droplet, $\Delta\mu$, to the (normalized) difference in solid-liquid step free energy between a zinc-blende and a wurtzite nucleus, $\Delta\gamma/\gamma_{\text{sl,ZB}}$, separates the domains of wurtzite and zinc-blende nanowire growth (see Supplementary Information 2). As elaborated in the Supplementary Information, we argue that the main effect of adding zinc during growth is a decrease in $\Delta\gamma/\gamma_{\text{sl,ZB}}$, which means a lowering of the liquid-solid step energy for zinc blende as compared to wurtzite. This suggests a strong interaction of the zinc atoms with the InP

¹Materials Innovation Institute (M2i), 2628CD Delft, The Netherlands. ²Philips Research Laboratories Eindhoven, High Tech Campus 11, 5656AE Eindhoven, The Netherlands. ³IMM, Solid State Chemistry, Radboud University Nijmegen, Heijendaalseweg 135, 6525AJ Nijmegen, The Netherlands. [†]Present address: Solid State Physics, Lund University, Box 118, S-221 00 Lund, Sweden.

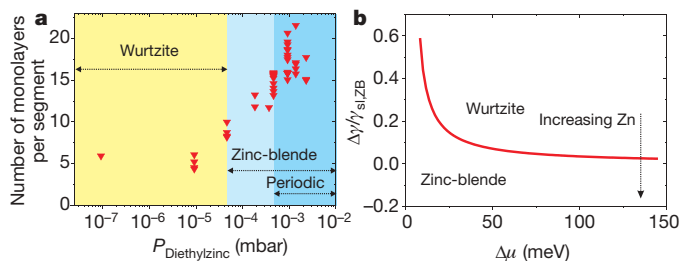


Figure 1 | The effect of Zn-doping on the InP nanowire crystal structure. **a**, The average segment length (number of monolayers, ML, per segment) between two adjacent stacking faults or twin planes as a function of the diethylzinc concentration. A transition from the wurtzite to the zinc-blende crystal structure is observed at 4.6×10^{-5} mbar diethylzinc, and above 4.6×10^{-4} mbar the twin planes are periodic for a given dopant level. The data refer to wires with diameters between 15 and 25 nm. **b**, Calculated kinetic phase diagram showing the domains of wurtzite and zinc blende as a function of $\Delta\gamma/\gamma_{sl,ZB}$ and $\Delta\mu$. Zn reduces the liquid–solid step energy of the zinc-blende structure with respect to the wurtzite structure and thus promotes the formation of zinc-blende-structured InP wires.

growth interface, as was also observed during electrical resistance measurements of Au(Zn)–InP contacts^{21,22}. In general, for nanowires grown by the VLS mechanism, the crystal structure may be intrinsically different from that of the bulk material and will depend on the combination of semiconductor with catalyst material.

Having adequate control over the nanowire crystal structure, we can now address the formation of TSLs. In Fig. 2, images from transmission electron microscopy (TEM) are shown of Zn-doped (9.2×10^{-4} mbar) InP nanowires with nominal diameters of 10, 20, 50 and 100 nm. It is clear from the overview images in Fig. 2a that the periodically twinned structure is general, although not all of the wires have the optimal orientation with respect to the electron beam. The segment length is uniform (Fig. 2b) throughout the wires. From the high resolution images in Fig. 2c, we observe that the periodic nanowires have the zinc-blende crystal structure with

{111} side facets, which are not parallel with respect to the long nanowire axis. Using the high resolution images, the number of monolayers between successive twin planes was counted, and the data are plotted in the histograms shown in Fig. 3a (here, a monolayer with thickness $d_{111} = 3.4$ Å contains pairs of In and P atoms). Segment lengths of 7 ± 2 , 13 ± 2 , 25 ± 3 and 33 ± 6 monolayers were found for wires with a diameter of 10, 20, 50 and 100 nm, respectively (Fig. 3a). Importantly, the periodicity in twinning is demonstrated by the relatively narrow distributions in segment lengths.

For the formation of the TSL, it is crucial that the {111}A and {111}B side facets are tilted in opposite directions (by $\theta = \theta_B = -\theta_A \approx 19.5^\circ$) with respect to the nanowire axis (Fig. 2c). As shown schematically in Fig. 4, at a certain moment during growth (situations 1 and 3) the top surface of the nanowire is a hexagon, and the shape of the catalyst droplet, connected to this surface, is close to spherical. When the wire grows, the {111}A edges move inward and their length increases, while the {111}B edges move outward and their length decreases. Thus the shape of the nanowire–droplet interface becomes increasingly triangle-like, as shown for situations 2 and 4 in Fig. 4b. This induces the catalyst droplet to distort so as to minimize its surface area, leaning over towards the long {111}A edges. At a certain point it becomes more favourable to form a twin plane and to start reducing the distortion of the catalyst particle by re-growth towards a hexagonal shape, rather than to continue growth towards a completely triangular shape. This mechanism of inverting triangularly shaped interfaces repeats itself continuously and produces the periodically structured wire.

To understand this process in more detail, we developed a quantitative model based on a specific mechanism relating the distortion of the droplet to the growth process (details are given in Supplementary Information 3). Nanowire growth proceeds layer by layer with a single nucleation event per layer, initiated at the nanowire edge¹². The free energy of formation of the nucleus depends on the difference between the liquid–solid contact angle and the tilt angle of the external facet of the nucleus. We have analysed this dependence quantitatively, making use of the simulator program Surface Evolver^{23,24}. It follows that the formation of nuclei with an external

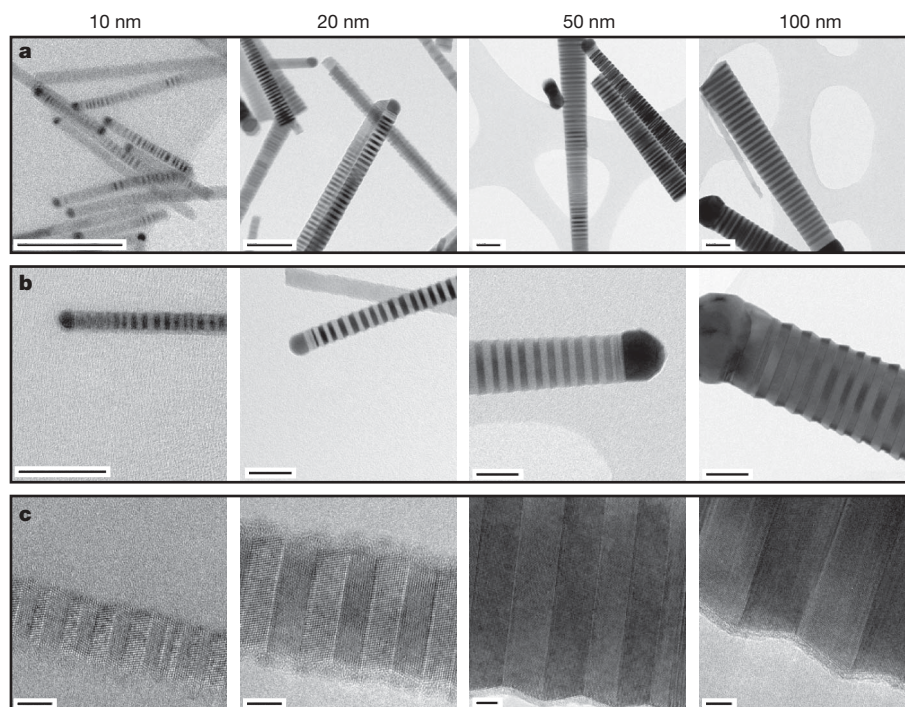


Figure 2 | TEM images of nanowire TSLs. **a–c**, Overview (**a**, **b**) and high resolution (**c**) TEM images of InP nanowires with a diameter of nominally 10, 20, 50 and 100 nm (columns 1–4, left to right) and a diethylzinc partial pressure of 9.7×10^{-4} mbar. Scale bars: **a**, 100 nm; **b**, 50 nm; and **c**, 5 nm.

The images are all taken close to the catalyst particle and therefore the observed side facets are due to the VLS growth. Lateral growth is also observed further down the nanowire.

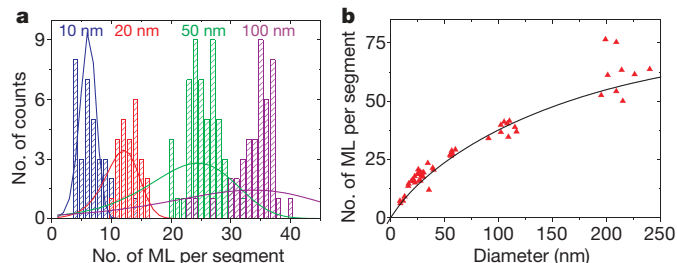


Figure 3 | The effect of wire diameter on the twin lattice spacing.

a, Histograms of the number of monolayers (ML) between two consecutive twin planes for wires with a nominal diameter of 10, 20, 50 and 100 nm. The narrow distributions demonstrate the periodicity. The curves give the calculated distributions obtained from the quantitative model, using the same parameter values as in **b**. **b**, The number of monolayers per segment versus the diameter. Each data point represents an averaged value of 25–50 segments, taken from a single nanowire with 9.7×10^{-4} mbar diethylzinc doping. The curve is the theoretical expression given in the text, with $\Delta = 4.5$ and $N_c = AD$ with $A = 0.35 \text{ nm}^{-1}$.

(solid–vapour) A-facet is strongly suppressed. As a result, the dominant processes are nucleation of external-B-facet nuclei at either B edges, adding another facet-conserving zinc-blende layer, or at A edges, introducing a layer that involves a twin plane and initiating re-growth. During facet-conserving growth, the liquid–solid contact angles increase (decrease) at the A edges (B edges) owing to the progressing droplet distortion illustrated in Fig. 4.

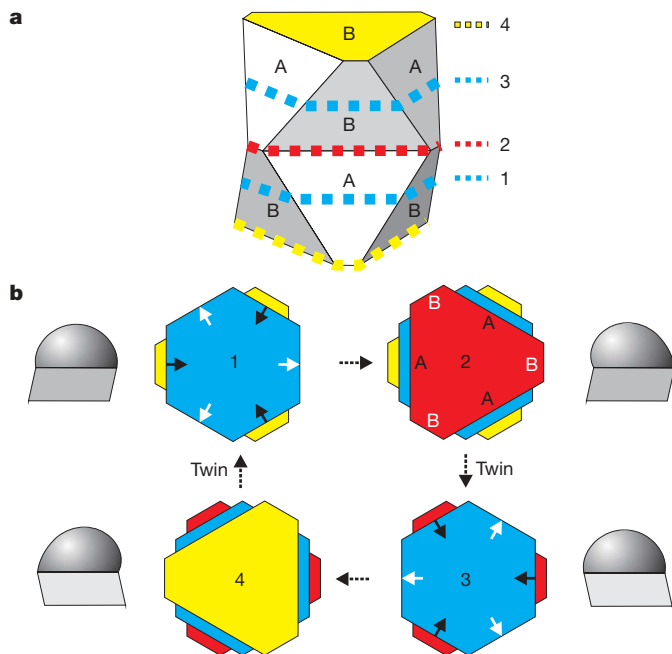


Figure 4 | Model for periodic twinning in nanowires. **a**, Schematic representation of the morphology of a twinned nanowire with the zinc-blende crystal structure with non-parallel {111} side facets. **b**, The cross-sectional shapes of the top facet of the nanowire crystal at the solid–liquid interface during growth. The numbers correspond to the positions indicated in **a**. Owing to the non-parallel orientation of the side facets, {111}A edges increase and {111}B edges decrease in length during vertical growth, and as a result a hexagonal interface develops into a triangle-like shape. At a certain moment, it is energetically more favourable to create a twin plane rather than to continue growing towards a fully triangular top interface. After twin formation, a triangle-like shape evolves back to a hexagonal shape and the cycle is repeated as schematized in **b**. To the left or right the corresponding calculated shape of the catalyst particle on a hexagonal (1 and 3) and a triangularly deformed (2 and 4) interface is depicted, showing the skewing of the droplet towards the long {111}A side edge and demonstrating that the contact angles depend sensitively on the cross-sectional shape.

We find that the contact angles depend linearly on the ratio of the wire height $H = Nh$ (with N the number of monolayers measured from the last hexagonal cross-section, and h the layer thickness) and the wire diameter D . This is similar to what has been observed for sawtooth faceting in Si nanowires and tentatively explained by considerations based on total energies instead of nucleation energies²⁵. From this, it follows that the difference in free energy of the two competing nucleation processes, $\delta\Delta G^*$, decreases linearly with wire height as $\delta\Delta G^* = \delta\Delta G_0^* - CNh/D$, where $\delta\Delta G_0^*$ is the free energy difference at hexagonal interface shape and C is a constant with the dimension of energy. As a result, we find that the critical number of monolayers, N_c , at which twin formation becomes the more favourable process, is proportional to the wire diameter D . However, the segment length will be substantially less than $2N_c$, because it is determined by the probability of an uninterrupted series of facet-conserving nucleations followed by a facet-changing nucleation. Taking this statistical aspect into account, we find that the number of layers in a segment is given by $N_s = 2N_c(1 + (1/\Delta)\ln[1 - \exp(-\Delta/N_c)])$, where $\Delta = \delta\Delta G_0^*/k_B T$ (here k_B is the Boltzmann constant and T is temperature). We find excellent agreement between this expression and the observed diameter dependence (Fig. 3b), if we make use of the explicit expressions (see Supplementary Information 3, and L.-F.F., manuscript in preparation) for $\delta\Delta G_0^*$ and N_c in terms of the physical parameters. These parameters are the solid–liquid surface tension γ_{SL} , liquid–vapour surface tension γ_{LV} , surface energy of a twin plane γ_T , supersaturation $\Delta\mu$, tilting angle θ , contact angle at hexagonal interface shape β_0 , and temperature T . In Fig. 3 we used $T = 713 \text{ K}$, $\gamma_T = 0.009 \text{ J m}^{-2}$ (refs 18, 26), $\gamma_{SL} \approx \gamma_{SV} \approx 0.8 \text{ J m}^{-2}$ (ref. 27), $\gamma_{LV} = 1.0 \text{ J m}^{-2}$ (in between the values for liquid Au and In), with $\beta_0 \approx 98^\circ$, and $\Delta\mu = 180 \text{ meV}$ per atom pair, which are physically plausible values.

In addition, the statistical analysis yields the width of the segment length distribution (Fig. 3a), using the same parameter values. The trend of increasing width with diameter is correctly reproduced, but the effect is overestimated for the thicker wires. This cannot solely be explained by the probably slightly smaller supersaturation for the larger-diameter nanowires, and could indicate that other energies or effects not included in the model might play a role here.

In general, to obtain a nanowire TSL with a narrow segment distribution, a large $\delta\Delta G_0^*$ and fairly small N_c are required. As $\delta\Delta G_0^*$ depends on supersaturation as $\delta\Delta G_0^* \propto (1/\Delta\mu)^2$, and N_c on supersaturation and wire diameter as $N_c \propto D/\Delta\mu$, it is clear that the supersaturation, determined by the trimethylindium flow, should be chosen low in order to increase $\delta\Delta G_0^*$, and the accompanying increase of N_c be compensated by having a small wire diameter, determined by the catalyst size.

Taking the above design rules into consideration, we have recently obtained a TSL structure in Zn-doped GaP nanowires, albeit with a broader segment length distribution. This preliminary result from a different materials system substantiates the general nature of the present approach.

Our insight into the formation of twins allows the fabrication of more complex structures by varying the Zn concentration during growth. Without Zn, random stacking faults in a wurtzite crystal structure should occur, and with Zn present the twinning should become periodic. This is indeed the case, as demonstrated by the TEM images in Fig. 5a and b, showing a wire for which a diethylzinc partial pressure of 9.2×10^{-4} mbar has been used intermittently during growth. The tapering of the nanowire is due to sidewall growth, which preferentially occurs on the zinc-blende sections. In Fig. 5c, the length of the zinc-blende sections has been plotted against the time interval during which the Zn precursor gas flow was switched on. The fitted linear curve has an insignificant offset from zero, suggesting an almost immediate switching from the wurtzite to the zinc-blende phase and vice versa. In Fig. 5d the number of monolayers between two twin planes is presented for the first four zinc-blende sections. The average segment length is 13 ± 3 monolayers,

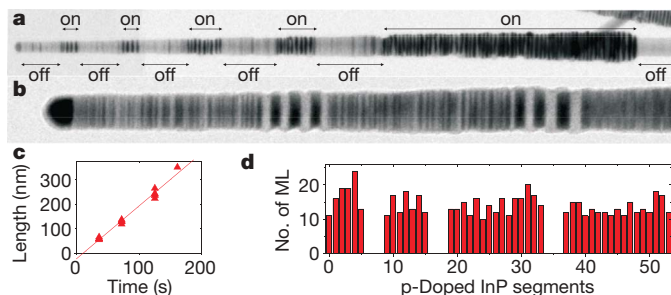


Figure 5 | InP nanowire with alternating periodic and non-periodic segments. **a**, An overview TEM image of a wire containing segments of intrinsic InP (containing randomly distributed stacking faults in a wurtzite structure) and Zn doped (9.7×10^{-4} mbar) segments (with different lengths) with periodic twin planes in a zinc-blende structure. The 'on' represents the p-doped segments where diethylzinc is added, whereas the 'off' represents the intrinsic, undoped, segments. **b**, Higher magnification TEM image of the segments closest to the gold particle. **c**, Zinc-blende section length versus the time interval during which the diethylzinc precursor gas flow was switched on. The intercept at approximately zero shows abrupt switching between the zinc-blende and wurtzite crystal structure. **d**, Number of monolayers between two successive twin planes in each p-doped segment obtained from HRTEM. An average of $\sim 13 \pm 3$ monolayers is found.

showing that the periodicity is clearly preserved in these short sections and is not affected by the growth history.

We have presented a viable and general route for the fabrication of TSLs in nanowires by controlling the nanowire morphology. This new instrument for manipulating the electronic properties of nanowires can be combined with already demonstrated features such as axial and radial heterostructures and doping profiles, further expanding the nanowire toolbox.

METHODS SUMMARY

The InP nanowires were synthesized in a low pressure (50 mbar) Aixtron 200 MOVPE reactor on InP (111)B substrates. The substrates were treated with a piranha etch for 1 min to remove the surface oxide, before deposition of Au colloids of different diameters (ranging from 10 to 200 nm). The nanowires were grown in the VLS growth mode using trimethylindium and phosphine as precursors, at partial pressures of 1.19×10^{-3} and 4.17×10^{-1} mbar, respectively, in a total flow of 61 min^{-1} hydrogen carrier gas. As dopant material diethylzinc was used for p-type doping. Before growth, an anneal step was carried out under phosphine/hydrogen atmosphere to desorb any surface oxide and to alloy the Au colloids with the InP substrate to ensure epitaxial growth. Growth was initiated when a temperature of 420°C was reached by switching on the trimethylindium. To change the dopant concentrations in the wires, the diethylzinc partial pressure was varied between 10^{-2} and 10^{-7} mbar, at constant trimethylindium and phosphine molar fractions in hydrogen. We point out that the molar fractions given for diethylzinc are those in the controlled gas flows in the reactor and not necessarily the built-in atomic fractions in the wires. The samples were analysed using TEM (FEI Tecnai 300 kV) in bright field, and high resolution TEM.

Received 8 April; accepted 16 October 2008.

- Gudikson, M. S. *et al.* Growth of nanowire superlattice structures for nanoscale photonics and electronics. *Nature* **415**, 617–620 (2002).
- Dick, K. A. *et al.* Synthesis of branched 'nanotrees' by controlled seeding of multiple branching events. *Nature Mater.* **3**, 380–384 (2004).
- Hochbaum, A. I. *et al.* Enhanced thermoelectric performance of rough silicon nanowires. *Nature* **451**, 163–168 (2008).

- van Dam, J. A. *et al.* Supercurrent reversal in quantum dots. *Nature* **442**, 667–670 (2006).
- Bao, J. *et al.* Optical properties of rotationally twinned InP nanowire heterostructures. *Nano Lett.* **8**, 836–841 (2008).
- Ikonic, Z. *et al.* Electronic properties of twin boundaries and twinning superlattices in diamond-type and zinc-blende-type semiconductors. *Phys. Rev. B* **48**, 17181–17193 (1993).
- Ikonic, Z. *et al.* Optical properties of twinning superlattices in diamond-type and zinc-blende-type semiconductors. *Phys. Rev. B* **52**, 14078–14085 (1995).
- Lui, A. S. *et al.* Advances in silicon photonic devices for silicon based optoelectronic applications. *Physica E* **35**, 223–228 (2006).
- Krames, M. R. *et al.* Status and future of high-power light-emitting diodes for solid-state lighting. *J. Display Tech.* **3**, 160–175 (2007).
- Mattila, M. *et al.* Crystal-structure-dependent photoluminescence from InP nanowires. *Nanotechnology* **17**, 1580–1583 (2006).
- Hiruma, K. *et al.* Growth and optical properties of nanometer scale GaAs and InAs whiskers. *J. Appl. Phys.* **77**, 447–462 (1995).
- Johansson, J. *et al.* Structural properties of $\langle 111 \rangle$ -B-oriented III–V nanowires. *Nature Mater.* **5**, 574–580 (2006).
- Xiong, Q. *et al.* Coherent twinning phenomena: Towards twinning superlattices in III–V semiconducting nanowires. *Nano Lett.* **6**, 2736–2742 (2006).
- Verheijen, M. A. *et al.* Three dimensional morphology of GaP–GaAs nanowires revealed by transmission electron microscopy tomography. *Nano Lett.* **7**, 3051–3055 (2007).
- Fissel, A. *et al.* Formation of twinning-superlattice regions by artificial stacking of Si layers. *J. Cryst. Growth* **290**, 392–397 (2006).
- Hibino, H. *et al.* Twinned epitaxial layers formed on Si(111) $\sqrt{3} \times \sqrt{3}$ -B. *J. Vac. Sci. Technol. A* **16**, 1934–1937 (1998).
- Hao, Y. *et al.* Periodically twinned nanowires and polytypic nanobelts of ZnS: The role of mass diffusion in vapor-liquid-solid growth. *Nano Lett.* **6**, 1650–1655 (2006).
- Akiyama, T. *et al.* An empirical potential approach to wurtzite-zinc-blende polytypism in group III–V semiconducting nanowires. *Jpn. J. Appl. Phys.* **45**, L275–L278 (2006).
- Glas, F. *et al.* Why does wurtzite form in nanowires of III–V zinc blende semiconductors? *Phys. Rev. Lett.* **99**, 146101 (2007).
- Minot, E. D. *et al.* Single quantum dot nanowire LEDs. *Nano Lett.* **7**, 367–371 (2007).
- Malina, V. *et al.* Effect of deposition parameters in the electrical and metallurgical properties of Au–Zn contacts to p-type InP. *Semicond. Sci. Technol.* **9**, 1523–1528 (1994).
- Weizer, G. W. *et al.* Au/Zn Contacts to p-InP: Electrical and Metallurgical Characteristics and the Relationship Between Them (NASA Technical Memorandum 106590, 1994).
- Brakke, K. E. The surface evolver. *Exp. Math.* **1**, 141–165 (1992).
- Brakke, K. E. The Surface Evolver Version 2.30. (<http://www.susqu.edu/brakke/evolver/evolver.html>) (2008).
- Ross, F. M. *et al.* Sawtooth faceting in silicon nanowires. *Phys. Rev. Lett.* **95**, 146104 (2005).
- Hurle, D. T. J. A mechanism for twin formation during Czochralski and encapsulated vertical Bridgman growth of III–V compound semiconductors. *J. Cryst. Growth* **147**, 239–250 (1995).
- Liu, Q. K. K. *et al.* Equilibrium shapes and energies of coherent strained InP islands. *Phys. Rev. B* **60**, 17008 (1999).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This research was carried out under project number MC3.05243 in the framework of the strategic research programme of the Materials Innovation Institute (M2i), the former Netherlands Institute of Metals Research, the FP6 NODE (015783) project, the Ministry of Economic Affairs in the Netherlands (NanoNed) and the European Marie Curie programme. We thank H. de Barse and F. Holthuysen for SEM imaging and P. van der Sluis, H. Wondergem and M. Decré for discussions.

Author Contributions All authors contributed to the design of experiments. G.I. was responsible for MOVPE growth, and M.A.V. for the TEM experiments. R.E.A. and M.A.V. analysed the TEM data. L.-F.F. and W.J.P.v.E. analysed the data quantitatively. R.E.A., L.-F.F., W.J.P.v.E. and E.P.A.M.B. co-wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to E.P.A.M.B. (erik.bakkers@philips.com).

Glacial greenhouse-gas fluctuations controlled by ocean circulation changes

Andreas Schmittner¹ & Eric D. Galbraith²

Earth's climate and the concentrations of the atmospheric greenhouse gases carbon dioxide (CO_2) and nitrous oxide (N_2O) varied strongly on millennial timescales during past glacial periods. Large and rapid warming events in Greenland and the North Atlantic were followed by more gradual cooling, and are highly correlated with fluctuations of N_2O as recorded in ice cores. Antarctic temperature variations, on the other hand, were smaller and more gradual, showed warming during the Greenland cold phase and cooling while the North Atlantic was warm, and were highly correlated with fluctuations in CO_2 . Abrupt changes in the Atlantic meridional overturning circulation (AMOC) have often been invoked to explain the physical characteristics of these Dansgaard–Oeschger climate oscillations^{1–3}, but the mechanisms for the greenhouse-gas variations and their linkage to the AMOC have remained unclear^{4–8}. Here we present simulations with a coupled model of glacial climate and biogeochemical cycles, forced only with changes in the AMOC. The model simultaneously reproduces characteristic features of the Dansgaard–Oeschger temperature, as well as CO_2 and N_2O fluctuations. Despite significant changes in the land carbon inventory, CO_2 variations on millennial

timescales are dominated by slow changes in the deep ocean inventory of biologically sequestered carbon and are correlated with Antarctic temperature and Southern Ocean stratification. In contrast, N_2O co-varies more rapidly with Greenland temperatures owing to fast adjustments of the thermocline oxygen budget. These results suggest that ocean circulation changes were the primary mechanism that drove glacial CO_2 and N_2O fluctuations on millennial timescales.

Our model simulates the coupled ocean–atmosphere–sea-ice–biosphere system⁹; it includes a dynamic terrestrial vegetation and carbon cycle model¹⁰, and a three-dimensional ocean general circulation model with ocean ecosystem dynamics and cycling of nitrogen, phosphorous, oxygen and carbon. The model uses a simple, energy-balance atmosphere and was run under glacial conditions (Methods). The model is forced by varying idealized freshwater perturbations to the North Atlantic (Fig. 1), mimicking surrounding ice-sheet fluctuations. We note that the simulations are idealized, in that the true forcing behind Dansgaard–Oeschger (D–O) cycles remains unknown. The model forcing is thus arbitrary and was chosen only to trigger AMOC variations.

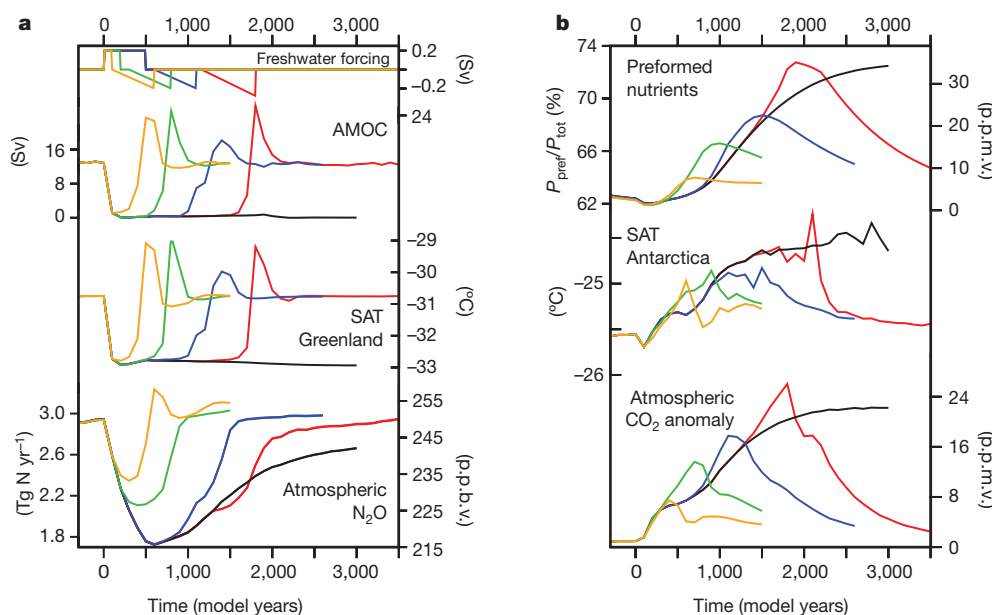


Figure 1 | Model simulations of glacial climate and greenhouse-gas fluctuations. The five model sensitivity runs differ in the length (yellow, 400 years; green, 700 years; blue, 1,100 years; and red, 1,700 years) of the simulated stadial (cold phase in Greenland). **a**, Time series of (from top) freshwater forcing, AMOC, Greenland surface air temperature (SAT), ocean

N_2O production (left scale) and atmospheric N_2O concentration (right scale). **b**, Time series of global fraction of preformed nutrients (left scale) and corresponding changes in atmospheric $p\text{CO}_2$ according to the simple theory²² (right scale), Antarctic SAT, and atmospheric CO_2 simulated by the complex model. Red lines correspond to the simulation shown in Figs 2 and 3.

¹College of Oceanic & Atmospheric Sciences, Oregon State University, Corvallis, Oregon 97331, USA. ²Atmospheric and Ocean Sciences, Princeton University, Princeton, New Jersey 08544, USA.

In response to the forcing, sinking of North Atlantic Deep Water (NADW) stops and the AMOC rapidly spins down from 13 Sv ($1 \text{ Sv} = 10^6 \text{ m}^3 \text{ s}^{-1}$) at model year 0 to almost 0 Sv after 100 years. Five sensitivity experiments have been conducted to assess the influence of the duration of the AMOC oscillations. In four experiments, the AMOC is switched back on after 400, 700, 1,100 and 1,700 years, respectively. In one experiment, it remains turned off (black lines, Fig. 1). The modelled climatic response, including rapid cooling in the North Atlantic and gradual warming in the Southern Hemisphere, is caused by reduced northward heat transport in the Atlantic, as described in detail elsewhere^{11,12}. It is qualitatively consistent with reconstructions but quantitatively underestimates the surface air temperature changes over Greenland and Antarctica, presumably owing to missing atmospheric dynamics. The experiments were not designed to reproduce any particular observed event. We choose the time period from 50 to 45 kyr before present (BP) for the comparison in Fig. 2 because the duration of the stadial phase corresponds well to one of our experiments.

After the AMOC collapse (following year 0), marine N_2O production rapidly decreases by 40% to less than 1.8 Tg N yr^{-1} during year 600 (Fig. 1). Figure 3 shows that N_2O production decreases almost

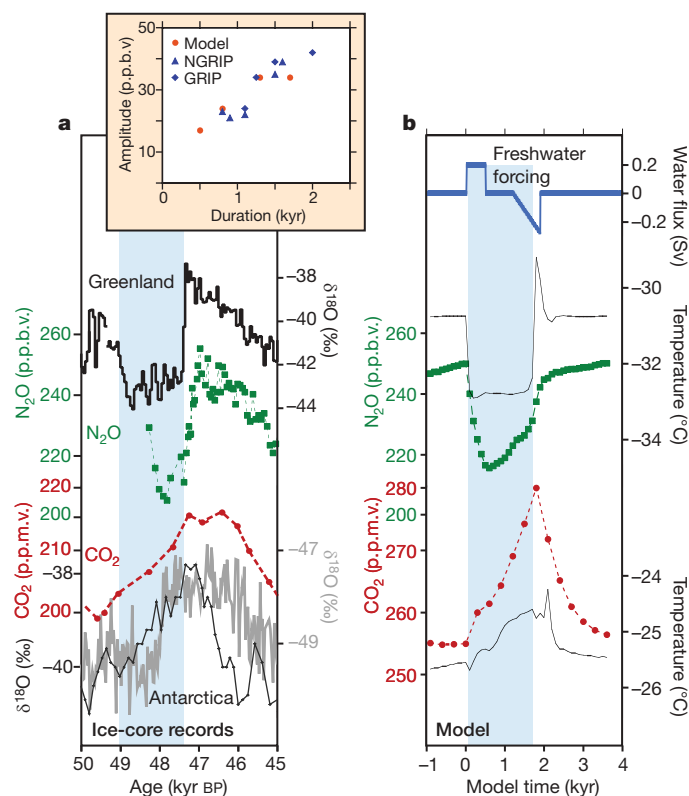


Figure 2 | Example of millennial changes in climate and greenhouse gases as recorded in polar ice cores from 45 to 50 kyr BP in comparison with model simulation. **a**, Ice-core records; **b**, model. Light blue shading indicates the stadial (cold in Greenland) phase following D-O interstadial 13 and immediately followed by rapid warming into D-O interstadial 12. This stadial coincides with an event of widespread ice-rafted debris in North Atlantic sediments (Heinrich event H5). In **a**, Greenland observations include $\delta^{18}\text{O}$ (ref. 28; temperature proxy) and N_2O concentrations¹⁵; Antarctic records include $\delta^{18}\text{O}$ (Dronning Maud Land²⁹, grey, and Byrd¹⁷, black), and CO_2 (Taylor Dome³⁰) on the age scale of ref. 24, shifted from the GISP2 age scale by 1.8 kyr to synchronize with the NGRIP age scale (Supplementary Information). In **b**, model output shows imposed North Atlantic freshwater forcing, temperature changes over Greenland (50° – 30° W, 68° – 78° N), atmospheric CO_2 and N_2O , and surface air temperature over Antarctica (68° – 78° S). Inset, amplitude of N_2O changes versus the duration of corresponding stadials, from two ice-core records following ref. 15 (blue), and from four model simulations (orange).

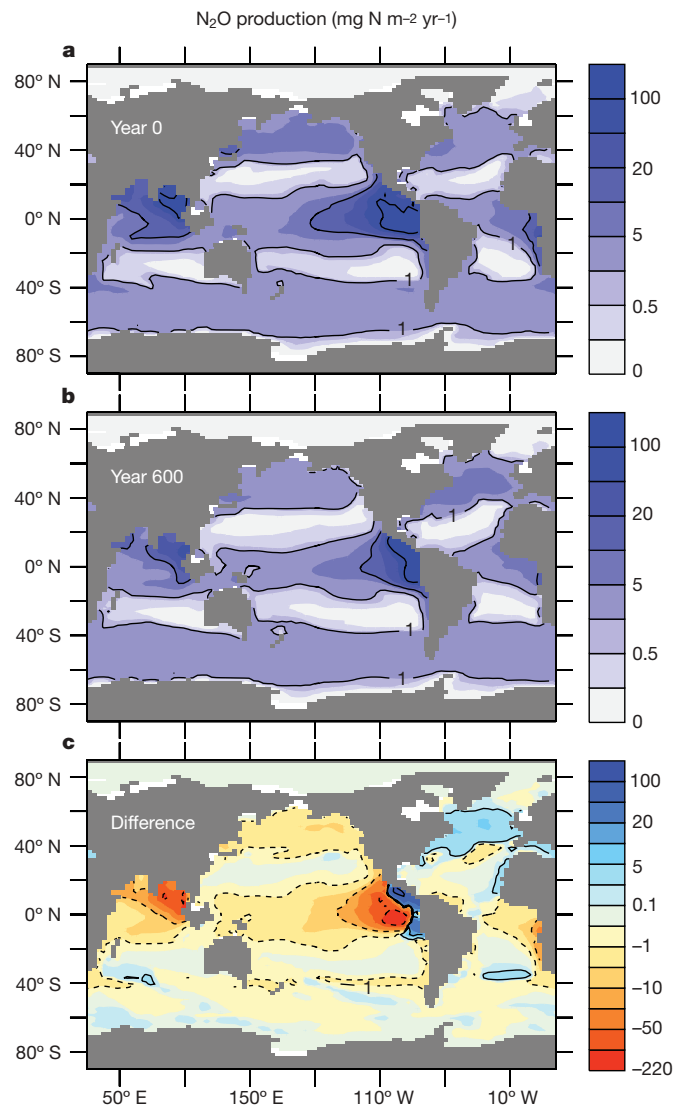


Figure 3 | Simulated ocean production of N_2O . **a**, During the interstadial at year 0; **b**, during the stadial at year 600; and **c**, the difference (year 600 minus year 0).

everywhere in the global ocean except for the North Atlantic. The decrease is largest in the low-oxygen regions of the eastern tropical Pacific and the northern Indian Ocean, but it is also noticeable in the North Pacific and elsewhere. Decreased productivity¹³ and better ventilation of thermocline waters lead to increasing subsurface oxygen concentrations, reducing Indo-Pacific N_2O production¹⁴. Simulated N_2O concentrations co-vary strongly with Greenland temperatures with little time lag (~ 100 years), consistent with the palaeorecord¹⁵. Modelled N_2O amplitudes are 15–40 parts per billion by volume (p.p.b.v.), and larger over longer stadials. This is in excellent agreement with ice-core data¹⁵ (Fig. 2 inset) and consistent with the earlier finding¹⁵ that N_2O increases are larger for longer interstadials, which generally follow longer stadials (see also Supplementary Note). After the rapid contemporaneous increases in simulated Greenland temperatures and N_2O concentrations, they remain constant, in contrast to observations, owing to the idealized forcing applied (Supplementary Fig. 3) and the lack of atmospheric and ice-sheet variability. Despite this difference, the simulated N_2O concentrations and Greenland temperatures remain highly correlated, consistent with the observations.

The amplitude of our simulated N_2O variability is much larger than that found in a previous study (10 p.p.b.v.) that used a zonally averaged ocean model⁸. As noted above, the simulated N_2O production changes show a strong zonal structure and, together with the

nonlinear dependence of N_2O production on oxygen concentrations, suggest that a zonally averaged model leads to a systematically biased underestimate. Sensitivity tests showed that the simulated amplitude is insensitive ($<5\%$) to parameter uncertainties within their 1σ range of the empirical N_2O production equation¹⁶, and that it is only moderately sensitive to the glacial ratio of marine to terrestrial N_2O production. Decreasing this ratio from 1/3 to 1/4 causes only a 15% reduction of the N_2O amplitude.

Methane variations on millennial timescales testify to changes in terrestrial systems during D-O events¹⁷, which presumably also altered N_2O production rates to some degree. However, the history of N_2O changes recorded in ice cores is markedly different from that of methane¹⁵, and the fact that our model can reproduce the correct amplitude of the glacial N_2O variations shows that changes in ocean sources could have dominated glacial variability of atmospheric N_2O on millennial timescales, obviating the need to invoke large changes in the terrestrial N_2O source^{15,18}.

The simulations also resolve a puzzling feature of the N_2O observations. Reference 15 notes that for long D-O oscillations, N_2O begins to increase before the rapid warming in Greenland, suggesting a lead and potentially a causal relationship for the AMOC resump-tions. The simulations show a comparable lead, but our experimental set-up precludes N_2O effects on climate and hence N_2O variations cannot influence the model AMOC. Rather, the recovery of N_2O is related to the long-term adjustment of the upper-ocean nitrate and oxygen inventories after the AMOC collapse. After the initial decrease, upper-ocean nutrient levels and global productivity climb after year 600 (Supplementary Fig. 3), followed closely by N_2O production. Most of this increase occurs in the North Atlantic and Arctic oceans, as nutrient inventories in these basins slowly adjust to the altered circulation.

In contrast to N_2O , simulated CO_2 concentrations increase slowly by about 25 p.p.m.v., on a millennial timescale, after the AMOC shuts down (Fig. 1), and decrease again only after the AMOC has resumed, co-varying strongly with Antarctic temperature and ventilation in the Southern Ocean¹⁹. Despite an increase in terrestrial carbon on a millennial timescale, atmospheric CO_2 increases owing to a larger-amplitude decrease in the marine carbon inventory (Supplementary Fig. 1). This occurs despite a small increase in the solubility of CO_2 in the oceans^{4,19}, through a much larger reduction in the global efficiency of the oceanic biological pump, a mechanism which can be most clearly understood through the change in the ocean preformed nutrient inventory²⁰. When surface waters in regions with high nutrient concentrations—principally the Southern Ocean—sink into the ocean interior, the capacity of those unutilized (preformed) nutrients to sequester carbon via the biological pump goes unrealized (Supplementary Fig. 2). In contrast, when nutrient-depleted waters from the subtropical Atlantic flow north to sink as NADW, they entrain relatively little preformed nutrients, thereby encouraging a higher efficiency of the biological pump²¹. A recently proposed theory quantitatively links changes in the atmospheric partial pressure of CO_2 (p_{CO_2} , in p.p.m.v.) to changes in the preformed fraction ($P_{\text{pref}}/P_{\text{tot}}$) of the global nutrient inventory²²: $\Delta p_{\text{CO}_2} = 312 \times \Delta P_{\text{pref}}/P_{\text{tot}}$ (see Methods).

The quantity $P_{\text{pref}}/P_{\text{tot}}$ is closely correlated with p_{CO_2} in our simulations (Fig. 1), increasing from 62% to more than 72% on a millennial timescale. The simple theory yields p_{CO_2} changes of ~ 30 p.p.m.v., somewhat overestimating the changes simulated by the complex model. The discrepancy arises mostly from simulated changes in land carbon, which decreases during the first 250 years, giving rise to the rapid initial increase in atmospheric CO_2 by 5 p.p.m.v., before reversing sign to gradually damp the long-term p_{CO_2} increase (Supplementary Fig. 1; see Supplementary Information for a discussion on uncertainties in land carbon simulations).

During the simulated shutdown of the AMOC, two factors conspire to cause the decreased efficiency of the biological pump, both of which relate to the volumetric contributions of water mass

end-members to the ocean interior. First, diminished input of low-preformed-nutrient NADW to the deep ocean causes the ocean interior to gradually become more dominated by the high-preformed-nutrient waters of the Southern Ocean. Second, weakened Southern Ocean stratification, caused by reduced input of salt to the deep waters via NADW, allows more rapid production of waters with high preformed nutrients¹⁹.

Modelled p_{CO_2} variations (Fig. 1) are ~ 25 p.p.m.v. for long D-O oscillations, consistent with ice-core data^{23,24}, and smaller for shorter oscillations, in agreement with previous results¹⁹. The simulation of small-amplitude p_{CO_2} changes during short D-O oscillations, which are likely to have been more influenced by terrestrial processes, cannot currently be evaluated because of the coarse time resolution of available p_{CO_2} data, but may be testable in the future as higher resolution data become available²³. However, the simulated decrease of p_{CO_2} after the abrupt warming in Greenland is clearly faster than observed (Fig. 2; see also ref. 23). This discrepancy points to processes not captured by the model, such as the impact of wind-shift-induced precipitation changes on vegetation and land carbon storage⁴ and/or ocean–sediment interactions, and indicates that more work needs to be done in order to fully reproduce the observed evolution of glacial CO_2 fluctuations on millennial timescales.

The simulations point to an important mechanistic contrast between the oceanic control on these two greenhouse gases (CO_2 and N_2O), explaining how contrasting atmospheric histories^{15,23} can be generated through a unified oceanic process. The partial pressure of CO_2 gradually increases after the interstadial–stadial transition owing to release of carbon from the intermediate and deep ocean, related to changes in the global effectiveness of the biological pump¹⁹. The stadial decrease of N_2O , on the other hand, is much faster because it is controlled by adjustments of upper-ocean oxygen cycling. Our results emphasize the role of ocean circulation and biogeochemical cycling for atmospheric greenhouse-gas concentrations. Given model projections of slowing AMOC²⁵, and suggestions that the current ocean sink for carbon is already decreasing²⁶ and that the ocean source of N_2O might increase in the future⁹, further progress in understanding ocean biogeochemical cycles will be required to refine the quantification of climate sensitivity to anthropogenic forcing.

METHODS SUMMARY

A two-dimensional (vertically averaged) energy-moisture-balance atmospheric model, including prescribed seasonally varying winds, provides a thermodynamically consistent solution for land–ocean surface conditions, without the computationally demanding requirements of a complex three-dimensional atmosphere. The model is well tested, and the ocean biogeochemical tracer distributions are in good agreement with observations when integrated under present-day conditions⁹. To improve comparison with the palaeorecord, we simulate aspects of the glacial climate by (1) prescribing land surface conditions in the presence of Northern Hemisphere (Laurentide and Fennoscandian) ice sheets²⁷ (that is, albedo, topography, no vegetation), and (2) increasing the outgoing long-wave radiation at the top of the atmosphere by 2.4 W m^{-2} (a simple approximation of reduced atmospheric greenhouse-gas concentrations). See Supplementary Information for details of the simulation of the glacial back-ground climate.

Marine N_2O production was calculated from simulated oxygen concentrations and oxygen consumption rates using an empirical formula (equation (8) of ref. 16) that reflects both nitrification and denitrification pathways of N_2O production. N_2O production increases nonlinearly with decreasing oxygen concentrations, with most production in the oxygen-depleted regions of the thermocline in the eastern tropical Pacific and the northern Indian Ocean (Fig. 3). Terrestrial N_2O production and the stratospheric sink were assumed constant. See Supplementary Information for more details of the N_2O calculation.

The remineralized nutrient concentration is given by $P_{\text{remi}} = \text{AOU} \times R_{\text{P:O}}$, where AOU is the apparent oxygen utilization ($\text{AOU} = \text{O}_2^{\text{sat}} - \text{O}_2$), O_2 is the dissolved oxygen concentration in sea water, O_2^{sat} is the temperature dependent oxygen saturation concentration, and $R_{\text{P:O}}$ is a (constant) phosphorus to oxygen ratio. The efficiency of the biological pump can be expressed as the biologically sequestered fraction of the total nutrient inventory, $P_{\text{remi}}/P_{\text{tot}} = 1 - P_{\text{pref}}/P_{\text{tot}}$.

Received 29 April; accepted 30 September 2008.

1. Broecker, W. S., Peteet, D. R. & Rind, D. Does the ocean-atmosphere system have more than one stable model of operation? *Nature* **315**, 21–26 (1985).
2. Rahmstorf, S. Ocean circulation and climate during the past 120,000 years. *Nature* **419**, 207–214 (2002).
3. Stocker, T. F. Past and future reorganizations in the climate system. *Quat. Sci. Rev.* **19**, 301–319 (2000).
4. Menviel, L., Timmermann, A., Mouchet, A. & Timm, O. Meridional reorganizations of marine and terrestrial productivity during Heinrich events. *Paleoceanography* **23**, PA1203, doi:10.1029/2007PA001445 (2008).
5. Kohler, P., Joos, F., Gerber, S. & Knutti, R. Simulated changes in vegetation distribution, land carbon storage, and atmospheric CO₂ in response to a collapse of the North Atlantic thermohaline circulation. *Clim. Dyn.* **25**, 689–708 (2005).
6. Marchal, O., Stocker, T. F. & Joos, F. Impact of oceanic reorganizations on the ocean carbon cycle and atmospheric carbon dioxide content. *Paleoceanography* **13**, 225–244 (1998).
7. Scholze, M., Knorr, W. & Heimann, M. Modelling terrestrial vegetation dynamics and carbon cycling for an abrupt climatic change event. *Holocene* **13**, 327–333 (2003).
8. Goldstein, B., Joos, F. & Stocker, T. F. A modeling study of oceanic nitrous oxide during the Younger Dryas cold period. *Geophys. Res. Lett.* **30**, 1092, doi:10.1029/2002gl01418 (2003).
9. Schmittner, A., Oschlies, A., Matthews, H. D. & Galbraith, E. D. Future changes in climate, ocean circulation, ecosystems and biogeochemical cycling simulated for a business-as-usual CO₂ emission scenario until year 4000 AD. *Glob. Biogeochem. Cycles* **22**, GB1013, doi:10.1029/2007GB002953 (2008).
10. Meissner, K. J., Weaver, A. J., Matthews, H. D. & Cox, P. M. The role of land surface dynamics in glacial inception: a study with the UVic Earth System Model. *Clim. Dyn.* **21**, 515–537 (2003).
11. Schmittner, A., Saenko, O. A. & Weaver, A. J. Coupling of the hemispheres in observations and simulations of glacial climate change. *Quat. Sci. Rev.* **22**, 659–671 (2003).
12. Crowley, T. J. North Atlantic deep water cools the southern hemisphere. *Paleoceanography* **7**, 489–497 (1992).
13. Schmittner, A. Decline of the marine ecosystem caused by a reduction in the Atlantic overturning circulation. *Nature* **434**, 628–633 (2005).
14. Schmittner, A. *et al.* Large fluctuations of dissolved oxygen in the Indian and Pacific oceans during Dansgaard-Oeschger oscillations caused by variations of North Atlantic Deep Water subduction. *Paleoceanography* **22**, PA3207, doi:10.1029/2006PA001384 (2007).
15. Fluckiger, J. *et al.* N₂O and CH₄ variations during the last glacial epoch: Insight into global processes. *Glob. Biogeochem. Cycles* **18**, GB1020, doi:10.1029/2003GB002122 (2004).
16. Nevison, C., Butler, J. H. & Elkins, J. W. Global distribution of N₂O and the Δ N₂O-AOU yield in the subsurface ocean. *Glob. Biogeochem. Cycles* **17**, 1119 (2003).
17. Blunier, T. & Brook, E. J. Timing of millennial-scale climate change in Antarctica and Greenland during the last glacial period. *Science* **291**, 109–112 (2001).
18. Sowers, T., Alley, R. B. & Jubenville, J. Ice core records of atmospheric N₂O covering the last 106,000 years. *Science* **301**, 945–948 (2003).
19. Schmittner, A., Brook, E. J. & Ahn, J. in *Ocean Circulation: Mechanisms and Impacts — Past and Future Changes of Meridional Overturning* (eds Schmittner, A., Chiang, J. C. H. & Hemming, S. R.) 209–246 (Vol. 173, Geophysical Monograph Series, American Geophysical Union, 2007).
20. Sigman, D. M. & Haug, G. H. in *Treatise On Geochemistry* Vol. 6 (eds Holland, H. D. & Turekian, K. K.) 491–528 (Elsevier Science, 2003).
21. Toggweiler, J. R. *et al.* Representation of the carbon cycle in box models and GCMs — 2. Organic pump. *Glob. Biogeochem. Cycles* **17**, 1026, doi:10.1029/2001GB001401 (2003).
22. Ito, T. & Follows, M. J. Preformed phosphate, soft tissue pump and atmospheric CO₂. *J. Mar. Res.* **63**, 813–839 (2005).
23. Ahn, J. & Brook, E. J. Atmospheric CO₂ and climate on millennial time scales during the last glacial period. *Science* **322**, 83–85 (2008).
24. Ahn, J. & Brook, E. J. Atmospheric CO₂ and climate from 65 to 30 ka BP. *Geophys. Res. Lett.* **34**, L10703, doi:10.1029/2007GL029551 (2007).
25. Schmittner, A., Latif, M. & Schneider, B. Model projections of the North Atlantic thermohaline circulation for the 21st century assessed by observations. *Geophys. Res. Lett.* **32**, L23710, doi:10.1029/2005GL024368 (2005).
26. Le Quere, C. *et al.* Saturation of the Southern Ocean CO₂ sink due to recent climate change. *Science* **316**, 1735–1738 (2007).
27. Peltier, W. R. Global glacial isostasy and the surface of the ice-age earth: The ICE-5G (VM2) model and GRACE. *Annu. Rev. Earth Planet. Sci.* **32**, 111–149 (2004).
28. Andersen, K. K. *et al.* High-resolution record of Northern Hemisphere climate extending into the last interglacial period. *Nature* **431**, 147–151 (2004).
29. EPICA. One-to-one coupling of glacial climate variability in Greenland and Antarctica. *Nature* **444**, 195–198 (2006).
30. Indermuhle, A. *et al.* Atmospheric CO₂ concentration from 60 to 20 kyr BP from the Taylor Dome ice core, Antarctica. *Geophys. Res. Lett.* **27**, 735–738 (2000).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This study was funded by the NSF Marine Geology and Geophysics programme grant 0728315-OCE, and the NSF Paleoclimate programme grant 0602395-ATM. We thank J. Ahn, E. Brook, R. Toggweiler and V. Brovkin for discussions and comments on the manuscript.

Author Contributions A.S. designed the study and performed the numerical experiments. E.D.G. performed the comparison with the proxy record. A.S. and E.D.G. discussed the results and contributed equally to the analysis of the data and writing of the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.S. (aschmitt@coas.oregonstate.edu).

Long-period earthquakes and co-eruptive dome inflation seen with particle image velocimetry

Jeffrey B. Johnson¹, Jonathan M. Lees², Alexander Gerst³, Dork Sahagian⁴ & Nick Varley⁵

Dome growth and explosive degassing are fundamental processes in the cycle of continental arc volcanism^{1,2}. Because both processes generate seismic energy, geophysical field studies of volcanic processes are often grounded in the interpretation of volcanic earthquakes³. Although previous seismic studies have provided important constraints on volcano dynamics^{4–6}, such inversion results do not uniquely constrain magma source dimension and material properties. Here we report combined optical geodetic and seismic observations that robustly constrain the sources of long-period volcanic earthquakes coincident with frequent explosive eruptions at the volcano Santiaguito, in Guatemala. The acceleration of dome deformation, extracted from high-resolution optical image processing, is shown to be associated with recorded long-period seismic sources and the frequency content of seismic signals measured across a broadband network. These earthquake sources are observed as abrupt subvertical surface displacements of the dome, in which 20–50-cm uplift originates at the central vent and propagates at $\sim 50 \text{ m s}^{-1}$ towards the 200-m-diameter periphery. Episodic shifts of the 20–80-m thick dome induce peak forces greater than 10^9 N and reflect surface manifestations of the volcanic long-period earthquakes, a broad class of volcano seismic activity that is poorly understood and observed at many volcanic centres worldwide⁷. On the basis of these observations, the abrupt mass shift of solidified domes, conduit magma or magma pads may play a part in generating long-period earthquakes at silicic volcanic systems.

Andesitic and dacitic domes, such as the recently active Soufrière Hills volcano (Montserrat)², Mount St Helens (USA)¹ or Galeras (Colombia)⁸, extrude lava over timescales of years to decades, but can also exhibit shorter-term variations in flux as well as periods of stagnancy and explosive self-destruction. Santiaguito's 86-year-long ongoing eruption⁹, for instance, has produced a 1-km³ edifice growing at a mean rate of $0.41 \text{ m}^3 \text{ s}^{-1}$ (ref. 10) with year-long surges approximately four times greater¹¹. Given these effusion rates and measured dome surface flow velocities, the near-vent flow thickness is estimated to be 20–80 m.

Although long-term fluctuations in dome effusion, controlled by magma supply and regional stresses¹², are quantifiable by means of remote sensing and field mapping, short-term variations, on the timescales of hours or less, are largely understood through numerical and theoretical modelling, which account for shallow degassing, slip-page and depressurization processes^{13,14}. Until recently, field corroboration of short-term eruptive phenomena have been lacking at silicic systems. New observations from Santiaguito highlight significant short-term dome emplacement processes coincident with earthquake generation.

We report here on episodic earthquakes recorded in conjunction with high-resolution video, revealing tens-of-centimetres uplift of a

crater-filling lava dome that is more than 30,000 m² in surface extent. These large and rapid mass shifts are coincident with explosive, ash-rich eruptions occurring 1–2 times per hour. Video image processing quantifies the shifted dome mass at $\sim 10^9$ – 10^{10} kg (assuming a thickness of 20–80 m), with local peak vertical accelerations of 2g–3g. Peak vertical momentum attained during rapid dome uplift is on the order of 10^9 kg m s^{-1} , with associated peak forces in the range $(0.1\text{--}3) \times 10^9 \text{ N}$ for various events.

Though Santiaguito dome shifts occur during a few seconds of each hour, they constitute much of the cumulative lateral flow (3–12 m per day) of the 200-m-diameter dome. Such rapid dome shifts, associated with the pyroclastic eruptions and earthquakes, have not previously been documented, but we propose that this is due more to the logistical difficulty of obtaining high-temporal-resolution geodetics of erupting domes than to an absence of this phenomenon at other silicic volcanoes. High-resolution videographic observations at Santiaguito demonstrate the usefulness of optical remote sensing for geodesy¹⁵ and, in particular, for monitoring abrupt Earth movement.

As one of the world's most reliably active dacite domes⁹, Santiaguito also possesses one of the best terrestrial viewing opportunities, with Santa Maria, Santiaguito's 'parent' stratovolcano, standing 1,200 m above and 2,500 m distant from the exploding and effusing Caliente dome (the current active vent). During a week of observation in January 2007, we deployed high-resolution video, a thermal imager and Doppler radar at the summit of Santa Maria to observe eruptions in tandem with a temporary six-station broadband seismometer and infrasound network. This network consisted of sites with 30-s and 120-s seismometers located 1–9 km from the active dome. Eruptions occurring ~ 40 times per day were accompanied by earthquakes, explosive infrasound pulses and rapid dome inflation recorded using high-resolution ($1,920 \times 1,080$ pixels) 30-fps video that focused on the extent of the 200-m-diameter dome.

Our videos show that explosive pyroclastic emissions coincided with systematic subvertical uplift of the lava dome surface. The extent and magnitude of uplift were quantified using particle image velocimetry (PIV) analysis, in which a sequence of co-eruption video stills is compared with a pre-eruption control image (see Supplementary Videos). Evolving surface dislocations orthogonal to the camera perspective were mapped for 50-m² sub-grids and indicate uplift initiation near the dome centre with progression out to the dome periphery over the course of a few seconds, until pyroclastic emissions obscured the dome surface, rendering PIV nonviable (Fig. 1).

Although three-dimensional dome motion is not uniquely derived from single-perspective videography, the inferred dome deformation is upward, as evidenced by a lack of downward-directed vectors beneath the central extrusion axis. The presumed vertical uplift rapidly approached an acceleration of more than $20\text{--}30 \text{ m s}^{-2}$ ($>2g$) in a few tenths of a second, which was sufficient to topple

¹Department of Earth and Environmental Science, New Mexico Tech, Socorro, New Mexico 87801, USA. ²Geological Sciences, University of North Carolina, Chapel Hill, North Carolina 27599, USA. ³Institute for Geophysics, University of Hamburg, Hamburg 20146, Germany. ⁴Earth and Environmental Sciences, Lehigh University, Bethlehem, Pennsylvania 18015, USA. ⁵Facultad de Ciencias, University of Colima, Colima 28045, Mexico.

meter-scale blocks on the dome surface. The instability of these blocks over the course of hours attested to regular and pronounced surface accelerations.

Rapid dome uplift began above the hypothesized feeding conduit, and detectable deformation propagated radially and reached the edge of the crater (>100 m distant) for the largest events. This observed

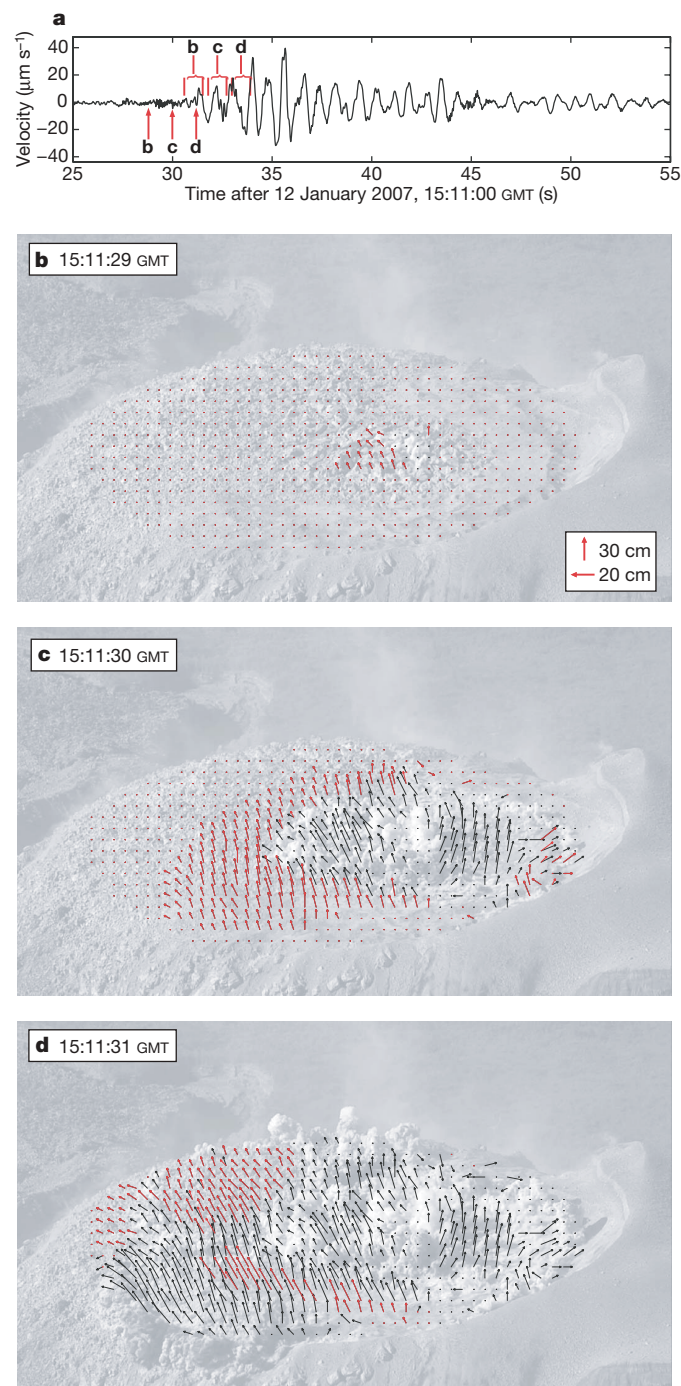


Figure 1 | Vertical seismogram and corresponding dome uplift for Santiaguito eruptive event. **a**, Seismogram. **b–d**, Sequential (1-s) images show dome dislocation vector field referenced to a pre-eruption image at 15:11:28 GMT. Red vectors indicate correlated movement in the indicated frame; black vectors show maximum dislocation. The vector scale is shown for pure horizontal and vertical motion. The vertical velocity seismogram was recorded at a station 3 km from the vent; the times corresponding to the image frames in **b–d** are indicated by arrows. Estimated transit times for P waves are 1.5–2.5 s and these corresponding time shifts are indicated by brackets for the three respective image frames.

deformation front propagated at $30\text{--}50\text{ m s}^{-1}$, which is too slow for typical elastic waves within solid or fluid magma, and too fast for buoyancy waves. More practically, the outward migration of dome inflation reflects incremental detachment and ‘flotation’ of a lava pad above a stationary basement. It is notable that large (25–100 m) ‘plates’ were often uplifted as intact units, inducing pyroclastic and gaseous emissions to concentrate around their peripheries through arcuate cracks noted by others^{16,17}. These fissures were ephemeral and evolved hourly as the dome effused and crack systems annealed or opened.

Dome uplift was coincident with the generation of long-period earthquakes, an important class of volcano earthquake (sometimes called B-type¹⁸) with a typical peak spectral energy of 0.5–2 Hz and poorly developed, frequently emergent body-wave arrivals. The sources of this broad earthquake class are poorly constrained because of strong anelastic attenuation and scattering of seismic waves in a typically heterogeneous volcanic edifice. Source mechanism interpretations are further complicated by the complexity of physical processes that occur within the volcano’s plumbing system^{7,19}.

At Santiaguito we directly observed surface manifestations of long-period events as conjoint subvertical uplift of the dome surface. Integrated surface motions, quantified using PIV, correlate favourably with seismic displacement trace amplitudes in the long-period band (Fig. 2). Taken together with temporal coincidence, this suggests that the primary long-period seismic phase is a response to momentum change imparted by the uplifted dome. The preponderance of upward-directed dislocation over an area that is small with respect to seismic long-period wavelengths ($\sim 10^3\text{ m}$) suggests that the seismic source can be approximated as a downward vertical point force, or recoil of the observed surface uplift. Volumetric displacement can be converted to integrated mass displacement, momentum records and force–time histories, which should in theory be linearly proportional to long-period displacements. We have calculated mass displacement histories assuming uplift of non-vesiculated lava (with uniform thickness of 20–80 m and density of $2,500\text{ kg m}^{-3}$) and computed vertical momentum and force histories as the first and second time derivatives of the uplift history, respectively (Fig. 3a–c). Under these assumptions, the events we observed using video impart downward-directed reactionary forces of $10^8\text{--}10^{9.5}\text{ N}$ to the volcanic edifice.

Amplitudes of single forces estimated from PIV are in general agreement with long-period waveforms recorded across the Santiaguito temporary seismic network. Twenty of the largest dome-shift earthquakes from 7 January 2007, for example, showed dynamic displacements in the long-period band that ranged from 7.5 to 23.5 μm at a station 1 km from the vent. We fitted these seismogram amplitudes and spectra with waveform synthetics calculated through the discrete wavenumber representation^{20,21}. Best-fit amplitudes of single forces ranged from 5×10^8 to $4.4 \times 10^9\text{ N}$, with pulse periods of 0.64 to 0.88 s (see Fig. 3d–f), which corroborate well with magnitudes estimated from PIV analysis and assumed dome thickness.

Vertical and subvertical single forces, including viscous conduit wall drag²², melt and gas fluid advection^{4,5}, and thrust response to pyroclastic emissions^{21,23–25} have previously been proposed as components, or the principal excitation mechanism, of the volcanic long-period elastodynamic response. We consider the possibility that Santiaguito’s recoil force, induced by rapid dome shift, is substantially greater than the hypothesized thrust contribution induced by low-density gas and ash ejection (see, for example, ref. 21). Venting gas, however, is probably responsible for concussion of the atmosphere and long-wavelength acoustic waves recorded by near-vent infrasound networks²⁶. The loose coupling of ground and atmospheric elastic waves, respectively induced by rapid dome uplift and volumetric gas expansion, explains how long-period seismic wave energy serves as a poor proxy for acoustic radiation and/or explosive eruption intensity (see, for example, ref. 27).

In our observations, long-period displacement spectral envelopes and optically inferred source force histories both peaked in the 0.5–2-

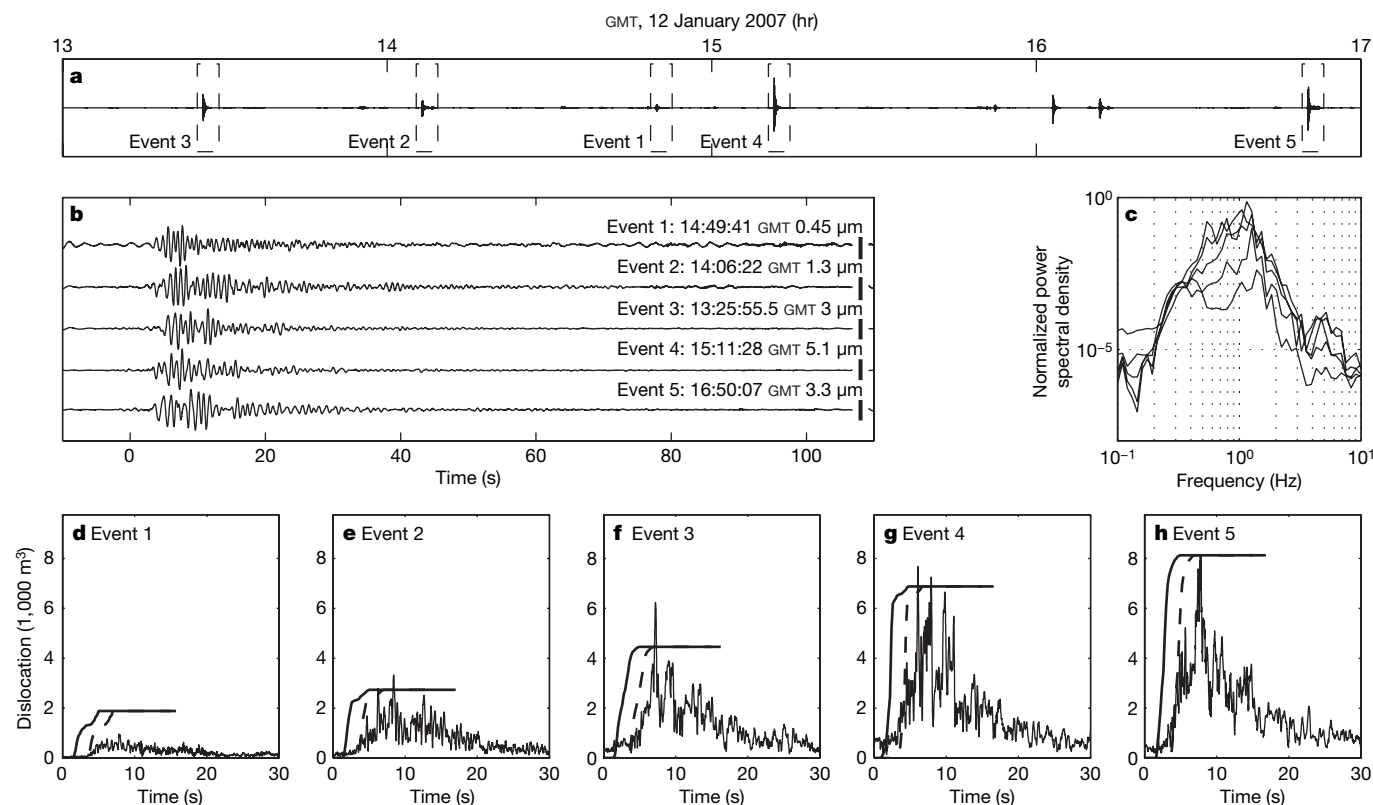


Figure 2 | Comparison of long-period earthquakes and dome motion.

a, Four-hour seismic record with five highlighted events of interest (**b**) and corresponding power spectral density normalized to event 4 (**c**). **d–h**, Time-synchronized, three-component seismic envelopes (jagged lines) and PIV-inferred cumulative volume dislocations (solid lines). Video animation of

these five events with vector displacement overlays are provided in Supplementary Information. Volume dislocation records (dashed lines) are time-shifted by expected seismic transit times and suggest coincident dome shift and elastic-wave radiation. Vertical seismic waveforms were recorded 3 km from the vent and have been filtered between 0.25 and 50 Hz.

Hz band, as expected for a linear convolution of a common source force function with the Earth's elastic response. Spectral consistency, temporal coincidence and correlation between dome shift and long-period magnitude strongly point to dome uplift as the probable

Santiaguito long-period source, and suggest that part of the mechanism for surficial long-period earthquakes at similar volcanoes worldwide might be attributable to rapid mass shifts. We note that although seismicity in the long-period band is the most energetic component of the recorded seismicity at many volcanoes, including Santiaguito, these events represent only a band-limited Earth response to a superposition of processes occurring within, and/or on the surface of, the volcanic edifice. Further discussion of broadband Santiaguito seismicity is provided in Supplementary Information.

Rapid dome uplift was accompanied by explosive degassing and the onset of lateral flow directed from the dome centre towards the crater periphery. Subsequent dome deflation was not traceable in the PIV analysis because of obfuscation by pyroclastic emissions. Santiaguito dome deflation, occurring less impulsively than inflation, was not evident in the long-period seismic record, but was suggested to be a 'slower' process by the simultaneously deployed Doppler radar system. Measured radio-wave phase changes indicated that a few centimetres (or more) surface deflation began only seconds after inflation onset.

We propose that each cycle of dome 'respiration'—upheaval followed by subsidence—is preceded and, consequently, triggered by steady accumulation of gas beneath the viscoelastic lava dome, which has 'annealed' and become impermeable during the 20–40-min

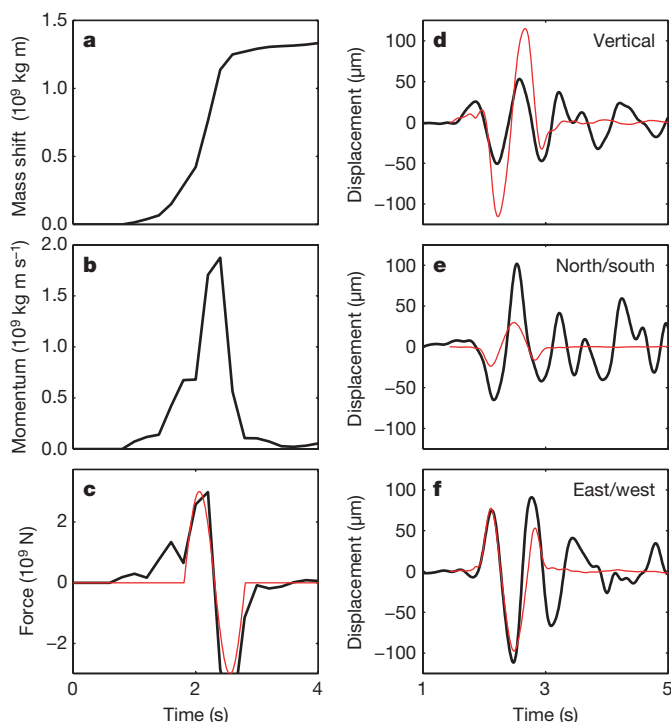


Figure 3 | PIV-derived vertical force and synthetic seismic waveform fit.

a–c, Source time history inferred from video recorded at 15:11:28 GMT (same event as in Fig. 1). Integrated vertical mass shift assuming uniform 80-m-thick dome (**a**); associated dome momentum history (**b**); vertical-force history (**c**), fitted by a sinusoid with a frequency of 1.2 Hz and an amplitude of $2.5 \times 10^9 \text{ N}$ (red). **d–f**, Three-component waveforms recorded 1 km from the vent (black) and synthesized waveforms (red) for modelled downward vertical force with time history opposite to that shown in c.

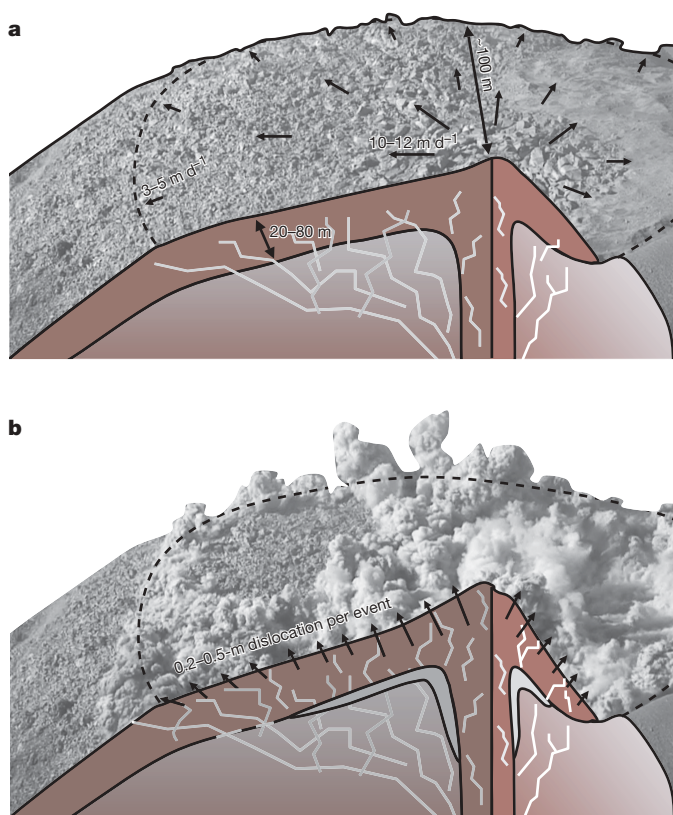


Figure 4 | Interpretation of episodic dome inflation. Cross-section showing inter-eruption recharge periods (occurring over tens of minutes; **a**) and sudden co-eruptive inflation events (with duration of several to tens of seconds; **b**). Surface radial vector flow rates in **a** are quantified from time-lapse footage recorded over four days. Uniform lava-pad thickness is estimated from these flow-field data and long-term effusion rates.

inter-eruption interval. Gas accumulation in fissures and/or voids beneath the lava dome, and magma conduit periphery, is fed by diffusion, disequilibrium degassing and heterogeneous flow¹³. These pressurizing reservoirs will not expand significantly until the normal tractions they exert, which are functions of pressure and the area of the gas/lava-pad surface contact, overcome the gravitational load and structural resistance of the overlying dome. At critical pressurization, the dome starts to detach and accelerate upward. The dome-averaged overpressure at the onset of flotation may be approximated as lithostatic pressures (5–20 bar for 20–80-m ‘carapace’ thicknesses) plus the driving pressures responsible for observed surface accelerations (0.03–1 bar on average beneath the entire dome). For some events, localized pressurization may exceed lithostatic pressure by a factor of two or more, judging from observations of $>20 \text{ m s}^{-2}$ localized accelerations.

Propagating detachment is speculatively associated with low-amplitude emergent broadband seismicity that is evident as long as 20 s before the primary long-period earthquake and coincident eruption (see Supplementary Information). As the mobile dome carapace separates from the basement, the pressurized surface area grows and the upward normal force lifting the bottom of the dome increases (Fig. 4). Progressing inflation provokes a positive feedback and rapid ‘buoying’ of the dome. In a few seconds, the 20–80-m-thick dome inflates by several tens of centimetres, resulting in strain rates of several per cent per second that induce brittle failure in portions of the lava carapace and facilitate explosive surface degassing. Infrasound coda lengths suggest that eruptions terminate within tens of seconds as degassing bleeds the gas reservoir pressure to below lithostatic levels.

Lateral dome flow coincident with uplift is evident as a component of dislocation in the PIV. Episodic lateral flow is presumably gravity

driven and facilitated by temporary decoupling of the lava dome from the underlying surface. A shallowly dipping (radially outward) contact is in agreement with the slightly bulged dome surface and inferred uniform dome thickness. This décollement presumably remains active owing to the exceptional lubrication present during episodic inflation.

METHODS SUMMARY

We used the discrete wavenumber representation^{20,21} to fit vertical force amplitudes with recorded seismic displacement waveforms at a station 1 km from the vent. To solve for carapace thickness, we minimized misfit between the synthetic and recorded waveforms for various near-surface sources. Sources were imbedded at the top of a 20-km-thick layered half-space, which exhibited increasing density ($1,800\text{--}2,400 \text{ kg m}^{-3}$), compressional-wave velocity ($2.8\text{--}4.7 \text{ km s}^{-1}$) and shear-wave velocity ($1.6\text{--}2.7 \text{ km s}^{-1}$). Recorded triaxial ground motions were fitted well by vertical single-force sinusoids and less well by other force combinations, such as isotropic and compensated linear vector dipoles. Corresponding vertical force source time histories were extracted from the PIV analysis, which is described more fully in Supplementary Information.

Received 12 June; accepted 11 September 2008.

- Iverson, R. M. *et al.* Dynamics of seismogenic volcanic extrusion at Mount St Helens in 2004–2005. *Nature* **444**, 439–441 (2006).
- Watts, R. B., Herd, R. A., Sparks, S. A. & Young, S. R. Growth patterns and emplacement of the andesitic lava dome at Soufrière Hills Volcano, Montserrat. *Mem. Geol. Soc. Lond.* **21**, 115–152 (2002).
- Young, S. R. *et al.* The ongoing eruption in Montserrat. *Science* **276**, 371–372 (1997).
- Waite, G. P., Chouet, B. A. & Dawson, P. B. Eruption dynamics of Mount St. Helens imaged from broadband seismic waveforms: Interaction of the shallow magmatic and hydrothermal systems. *J. Geophys. Res.* **113**, doi:10.1029/2007JB005259 (2008).
- Chouet, B., Dawson, P. B. & Arciniega-Ceballos, A. Source mechanism of Vulcanian degassing at Popocatepetl Volcano, Mexico, determined from waveform inversion of very long period signals. *J. Geophys. Res.* **110**, doi:10.1029/2004JB003524 (2005).
- Kanamori, H. & Given, J. W. Analysis of long-period seismic waves excited by the May 18, 1980 eruption of Mount St. Helens: a terrestrial monopole? *J. Geophys. Res.* **87**, 5422–5432 (1982).
- Chouet, B. A. Long-period volcano seismicity: its source and use in eruption forecasting. *Nature* **380**, 309–316 (1996).
- Gil Cruz, F. & Chouet, B. A. Long-period events, the most characteristic seismicity accompanying the emplacement and extrusion of a lava dome in Galeras Volcano, Colombia, in 1991. *J. Volcanol. Geotherm. Res.* **77**, 121–158 (1997).
- Rose, W. I. Volcanic activity at Santiaguito Volcano, 1976–1984. *Spec. Pap. Geol. Soc. Am.* **212**, 17–27 (1987).
- Anderson, S. W., Fink, J. & Rose, W. I. Mount St. Helens and Santiaguito lava domes: The effect of short term eruption rate on surface texture and degassing processes. *J. Volcanol. Geotherm. Res.* **69**, 105–116 (1995).
- Harris, A. J. L., Rose, W. I. & Flynn, L. Temporal trends in lava dome extrusion at Santiaguito, 1922–2000. *Bull. Volcanol.* **65**, 77–89 (2003).
- Manga, M. & Brodsky, E. E. Seismic triggering of eruptions in the far field: Volcanoes and geysers. *Annu. Rev. Earth Planet. Sci.* **34**, 263–291 (2006).
- Sahagian, D. Volcanic eruption mechanisms: Insights from intercomparison of models of conduit processes. *J. Volcanol. Geotherm. Res.* **143**, 1–15 (2005).
- Starostin, A., Barmin, A. & Melnik, O. A transient model for explosive and phreatomagmatic eruptions. *J. Volcanol. Geotherm. Res.* **143**, 133–151 (2005).
- Leprince, S., Berthier, E., Ayoub, F., Delacourt, C. & Avouac, J.-P. Monitoring earth surface dynamics with optical imagery. *Eos* **89**, doi:10.1029/2008EO10001 (2008).
- Bluth, G. J. S. & Rose, W. I. Observations of eruptive activity at Santiaguito volcano, Guatemala. *J. Volcanol. Geotherm. Res.* **136**, 297–302 (2004).
- Sahetapy-Engel, S. & Harris, A. J. L. Thermal structure and heat loss at the summit crater of an active lava dome. *Bull. Volcanol.* doi:10.1007/s00445-00008-00204-00443 (2008).
- McNutt, S. R. in *Monitoring and Mitigation of Volcano Hazards* (eds Scarpa, R. & Tilling, R.) 100–146 (Springer, 1996).
- Jousset, P., Neuberg, J. & Jolly, A. Modelling low-frequency volcanic earthquakes in a viscoelastic medium with topography. *Geophys. J. Int.* **159**, 776–802 (2004).
- Bouchon, M. Discrete wave number representation of elastic wave fields in three-space dimensions. *J. Geophys. Res.* **84**, 3609–3614 (1979).
- Nishimura, T. Source parameters of the volcanic-eruption earthquakes at Mount Tokachi, Hokkaido, Japan and a magma ascending model. *J. Geophys. Res.* **100**, 12465–12473 (1995).
- Ohminato, T., Chouet, B. A., Dawson, P. B. & Kedar, S. Waveform inversion of very-long-period impulsive signals associated with magmatic injections beneath Kilauea Volcano, Hawaii. *J. Geophys. Res.* **103**, 23839–23862 (1998).

23. Kanamori, H., Given, J. W. & Lay, T. Analysis of seismic body waves excited by the Mount St. Helens eruption of May 18, 1980. *J. Geophys. Res.* **89**, 1856–1866 (1984).
24. Chouet, B. *et al.* Source and path effects in the wave fields of tremor and explosions at Stromboli volcano. *J. Geophys. Res.* **102**, 15129–15150 (1997).
25. Brodsky, E. E., Kanamori, H. & Sturtevant, B. A seismically constrained mass discharge rate for the initiation of the May 18, 1980 Mount St. Helens eruption. *J. Geophys. Res.* **104**, 29387–29400 (1999).
26. Johnson, J. B., Aster, R. C. & Kyle, P. R. Volcanic eruptions observed with infrasound. *Geophys. Res. Lett.* **31**, doi:10.1029/2004GL020020 (2004).
27. Mori, J. *et al.* Seismicity associated with eruptive activity at Langila Volcano, Papua New Guinea. *J. Volcanol. Geotherm. Res.* **38**, 243–255 (1989).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the many personnel involved with instrumentation preparation and field work, including R. Sanderson, O. Marcillo, M. Hort, E. Sanchez, W. Rose, M. Ruiz, R. Salguero and staff at IRIS PASSCAL and INSIVUMEH (Guatemala). This project was made possible by the US National Science Foundation EAR grants 0440225 and 0440054.

Author Contributions J.B.J. performed the PIV and seismic analysis and wrote the paper with assistance from D.S. J.M.L. carried out the synthetic waveform modelling. A.G. and N.V. contributed to model development and data collection and analysis.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.B.J. (jeff.johnson@ees.nmt.edu).

LETTERS

Acoel development indicates the independent evolution of the bilaterian mouth and anus

Andreas Hejnol¹ & Mark Q. Martindale¹

Most bilaterian animals possess a through gut with a separate mouth and anus¹. It is commonly believed that during the transition from radial to bilateral symmetry, both openings evolved simultaneously by the lateral closure of a slit-like blastopore^{1–4}. Molecular phylogenies however, place the acoel flatworms, which have only one opening to their digestive system, as the sister group to all remaining Bilateria^{5–7}. To address how this single body opening is related to the mouth and anus of the protostomes and deuterostomes, we studied the expression of genes involved in bilaterian foregut and hindgut patterning during the development of the acoel *Convolutriloba longifissura*. Here we show that the genes *brachyury* and *gooseoid* are expressed in association with the acoel mouth, suggesting that this single opening is homologous to the mouth of other bilaterians⁸. In addition, we find that the genes *caudal*, *orthopedia* and *brachyury*—which are expressed in various bilaterian hindguts^{8–10}—are expressed in a small region at the posterior end of the animal, separated from the anterior oral *brachyury*-expressing region by a dorsal domain of ectodermal *bmp2/4* expression. These results contradict the hypothesis that the bilaterian mouth and anus evolved simultaneously from a common blastoporal opening, and suggest that a through gut might have evolved independently in different animal lineages.

Bilaterian animals are subdivided into two main clades, which were named according to the position of the mouth in relation to a transient embryonic structure called the blastopore¹¹. In deuterostomes ('secondary mouth') the site of gastrulation becomes the anus and the mouth is formed at a different site. However, in protostomes ('first mouth') the relationship between the site of gastrulation (the site at which the endomesoderm forms) and the site of mouth formation is much more variable than the clade name suggests¹. The evolutionary origin of the bilaterian mouth and anus, and their relationship to the blastopore, however, have an integral role in hypotheses about the evolution of bilateral animals from a radially symmetric ancestor with a single opening to its digestive cavity^{1–4,12–15}. Most models begin with either a radially symmetric larva or a cnidarian polyp-like organism that elongated its body along the future anterior–posterior axis, followed by a lateral closure of this slit-like blastoporal opening with the ends giving rise to a mouth and anus (Fig. 1a). Derivatives of this idea, which assumes the simultaneous evolution of mouth and anus, are the 'trochaea', the 'enterocoely', the 'bilaterogastraea'¹⁵ and the 'amphistomy' hypotheses². In contrast to these ideas is the suggestion that a posterior blastoporal opening of a radially symmetric planula-like organism shifted anteriorly, resulting in a bilateral organism with a single ventral 'mouth' (acoeloid-planuloid-hypothesis^{12,14,16}; Fig. 1a). Morphology¹⁷ and recent molecular phylogenies place the Acoela and the nemertodermatid flatworms as basal branches of the Bilateria^{5–7} (Fig. 1b). These bilaterally symmetric triploblastic worms possess a mid-ventral mouth and lack an anus, which suggests that a single opening to

the endodermal digestive system represents the ancestral state of the Bilateria (Fig. 1c–h). To investigate how the acoel mouth opening is related to the mouth and anus of protostomes and deuterostomes, we investigated the embryonic and juvenile expression of the genes that have been used to argue for the homology of the protostome and deuterostome foregut (oral ectoderm)⁸. The *brachyury* orthologue (*bra*) from the acoel *C. longifissura* contains the T-box DNA-binding site but, like cnidarians and ctenophores, it lacks the amino-terminal Smad-binding domain, which is found in most bilaterian orthologues¹⁸. *bra* is expressed in a domain (Fig. 2a) that gives rise to the ectoderm, anterior to and surrounding the juvenile mouth (Fig. 2b, c). However, in addition to the oral domain, we find a second domain of *bra* expression at later stages in the ectoderm at the posterior tip of the juvenile tail (Fig. 2b, c). This domain is located in the posterior ectodermal region that gives rise to the future gonopore (Fig. 2s–u) and follows the same temporal delay as the hindgut expression of *bra* in protostomes and deuterostomes^{19,20}. The *gooseoid* orthologue (*gsc*) in *C. longifissura* is expressed transiently in endomesodermal cells of the early embryo (Fig. 2d), but is downregulated during development, mirroring the localization and timing of the early function in endomesoderm specification found in some bilaterians²¹. Later in development, *gsc* is only expressed in the same oral ectodermal area as *bra* (Fig. 2e, f) but never in any posterior domain. Because *C. longifissura* lacks an internalized ectodermal component of the gut (foregut), the surface ectodermal expression of *gsc* and *bra* associated with the mouth supports the homology of the single body opening of acoels with the mouth of protostomes and ambulacrarians (echinoderms and hemichordates)⁸.

To understand the importance of the posterior expression domain of *bra*, we studied the expression of further genes expressed in the hindgut of some bilaterians, *caudal* (*cdx*), *orthopedia* (*otp*), *forkhead* (*FoxA*, also known as *HNF3*) and *NK2.1* (also known as *ceh-22*; see Supplementary Table 2). The gene *cdx* is primarily expressed in the deuterostome and protostome hindgut¹⁰. *cdx* is expressed in juvenile acoels in neural and mesodermal tissue along the entire anterior–posterior axis (Fig. 2g–i), reminiscent of the pattern seen in the polychaete annelid *Capitella*²². However, in addition to this broad expression domain, *cdx* is also expressed in the most posterior ectodermal tip of the animal, at the same position as *bra* expression (Fig. 2h, i). The gene *otp* is expressed in anterior neural tissue in many bilaterians, however, it is also expressed widely in the posterior hindgut of some protostomes and deuterostomes (except chordates)⁹. *otp* is expressed in the hindgut of *Drosophila*⁹ and in an internal posterior domain close to the future anus in the hemichordate *Saccoglossus kowalevskii*²³. In *C. longifissura*, *otp* expression appears in neural precursors of the brain commissure, in presumptive sensory cells of the ventral fold (Fig. 2j), and in a posterior ectodermal domain just internal to the location of *cdx* and *bra* expression (Fig. 2k, l). *FoxA* is transiently expressed in the endomesoderm after

¹Kewalo Marine Laboratory, PBRC, University of Hawaii, 41 Ahui Street, Honolulu, Hawaii 96813, USA.

gastrulation (Fig. 2m) and later in the endoderm bordering the mouth (Fig. 2n, o), but never in the posterior region. Orthologues of the *NK2.1* gene are expressed in diverse tissues and regions in bilaterians, for example, the apical organ, nerve cords, the thyroid gland and in the foregut and hindgut (Supplementary Table 2). In *C. longifissura*, *NK2.1* is expressed in putative neural cells (Fig. 2p) and later in the anterior nervous system (Fig. 2p, q) as well as in an expression domain that extends ventral-posterior to the expression domain of *bra*, *gsc* and *otp* (Fig. 2p, q). The similarity in the positional and tissue specificity of expression of these genes in acoels and bilaterians is notable. The localization of *NK2.1* (in the posterior ectoderm) and *FoxA* (in the endoderm at the border of the mouth) expression makes them ideal candidates for co-option in foregut and hindgut patterning in other bilaterians (for example, *FoxA* orthologues are expressed in both the foregut and endodermal midgut of several protostomes and deuterostomes; Supplementary Table 2). Phylogenetic analysis suggests that the endodermal expression of *FoxA* in acoels is ancestral, and demonstrates the co-option of *FoxA* for foregut patterning (Supplementary Table 2).

Chordates seem to express a unique complement of genes in the oral region that are not found in either the oral or anal openings of ambulacrarians and protostomes²⁴. We investigated the expression of the 'chordate mouth' genes, *six3/6*, *pitx* and *dlx* in *C. longifissura*, but none of these genes was expressed in association with either the acoel mouth or the posterior domain (Supplementary Fig. 1).

These findings further support the hypothesis that the chordate mouth either evolved secondarily or became shifted in relation to the mouth of the deuterostome ancestor^{2,24}.

To determine whether there is an embryonic relationship between the formation of the mouth and the two expression domains of *bra*, we compared our results to the embryonic fate map of acoel flatworms. Acoels show a stereotyped cleavage program known as 'duet' cleavage, and the embryonic fate map of *Neochildia fusca* indicates that the mouth is formed at the border of the micromeres 1b, 2a, 2b, 3a and 3b at the anterior-ventral edge of the site of gastrulation (Supplementary Fig. 2a)²⁵. In the juvenile, all the descendants of micromere 1b seem to co-express *bra* and *gsc* (compare Fig. 2b, c, e, f with Supplementary Fig. 2a). The acoel fate map also shows that the region of *bra*, *cdx* and *otp* expression at the posterior of the animal is formed by cells located at the exact opposite side of the blastopore from cells which form the mouth, that is, the border of the micromeres 1a, 2a, 2b, 3a and 3b. (Supplementary Fig. 2a)²⁵. The area between the two *bra* expression domains at the mouth and the posterior tip of the animal is formed by descendants of the posterior-most ectodermal micromeres (3a and 3b). The acoel fate map also allows us to identify that the earliest *bra*-expressing cells at the 16-cell stage are the micromeres 1a²² and 1b²² (Supplementary Fig. 2b–e). These micromeres are the posterior descendants of the animal-most pair of micromeres 1a and 1b. After the complete closure of the blastopore (Supplementary Fig. 3 and Supplementary Movie 1),

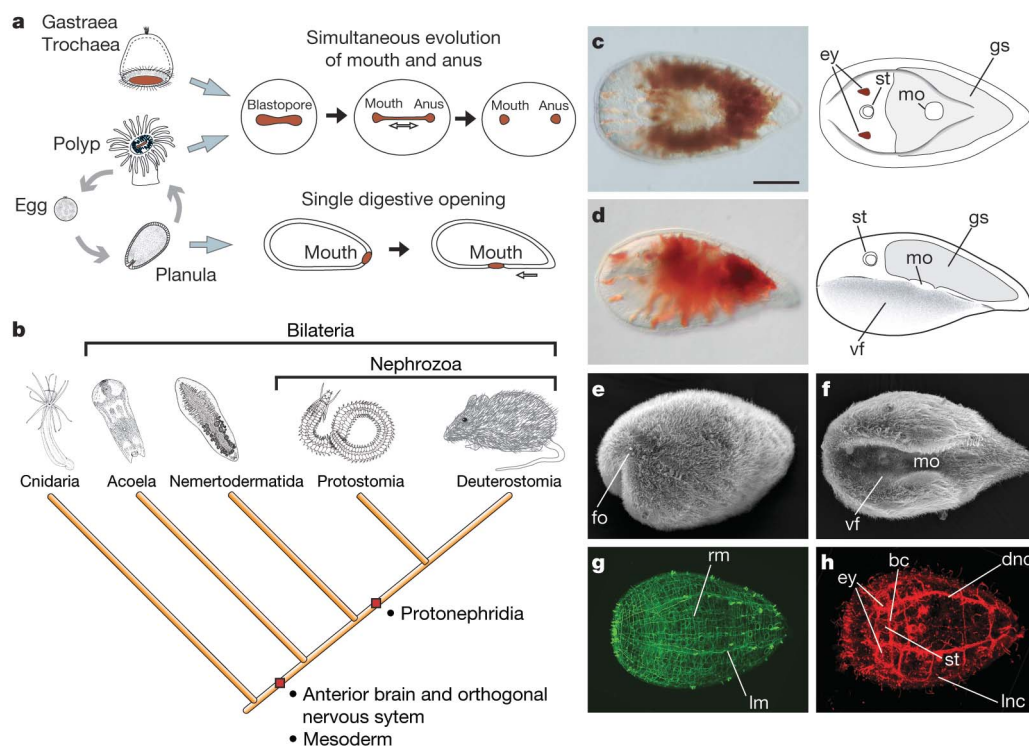


Figure 1 | Competing hypotheses for the evolution of the through gut and the position of acoel flatworms. **a**, Hypotheses based on a simultaneous evolution of mouth and anus (top) either start with a pelagic hypothetical ancestor (gastraea^{2,15} and trochaea¹) or start with a sessile cnidarian polyp-like ancestor³. The common assumption is that the single opening, arising from the blastopore, gave rise to mouth and anus simultaneously by a lateral closure. Alternative hypotheses (bottom) suggest a cnidarian planula-like ancestor with a single opening, which shifts ventrally during the transition to bilateral symmetry with an anus forming secondarily^{14,16}. **b**, Phylogenetic position of Acoela in relation to Cnidaria and Bilateria on the basis of recent molecular phylogenies^{5–7}. Red squares indicate evolutionary novelties

(apomorphies). **c**, **d**, Images and body plans of a juvenile of the acoel *C. longifissura* from dorsal (**c**) and lateral (**d**) views. **e**, **f**, Scanning electron microscopy images of a juvenile from lateral (**e**) and ventral (**f**) views. The mouth opening is located in the ventral fold (vf; **f**). **g**, Muscle net of the juvenile, labelled with phalloidin. The dorsal view is shown. **h**, Juvenile nervous system immuno-labelled with anti-tyrosinated tubulin antibody. The dorsal view is shown. The anterior is to the left for all images. bc, brain commissure; dnc, dorsal nerve cord; ey, eye; fo, frontal organ; gs, gut syncytium; lm, longitudinal musculature; lnc, lateral nerve cord; mo, mouth; rm, ring musculature; st, statocyst; vf, ventral fold. Scale bar, 100 μ m.

expression persists only in descendants of one micromere (1b), which contributes to the area that is anterior to the mouth (Supplementary Fig. 2a), whereas expression wanes in the 1a lineage until juvenile stages. The ectodermal region on the dorsal surface is generated by the 1a micromere (Supplementary Fig. 2f, g), whose asymmetric, extended posterior growth over the dorsal surface pushes tissue anteriorly—similar to the dorsal expansion of micromere descendants in spiralian taxa²⁵. These descendants of the 1a micromere, which form the dorsal ectoderm opposite the mouth site, express the orthologue of BMP2/4 (known as *dpp* in *Drosophila*; Supplementary Fig. 2f, g); BMP2 is responsible for dorsal patterning in protostomes and ambulacrarians^{26,27} suggesting a similar function in acoels.

Molecular and developmental cell lineage data suggest that the acoel mouth opening is homologous to the mouth of protostomes and deuterostomes and that the last common ancestor of the Bilateria (the 'urbilaterian'), had only this single digestive opening (Fig. 3). The existence of a posterior domain of coordinated expression of 'hindgut' genes in the acoel *C. longifissura*, however, can be explained in two different ways. A through gut could have been present in the last common ancestor of all bilaterians and the anus could have been lost independently in both Acoela and Nemertodermatida lineages. The expression of hindgut markers at the posterior pole in *C. longifissura* would therefore be remnants of a posterior anal opening. The more parsimonious explanation, however, is to assume that the metazoan mouth evolved first and the anal opening arose independently through co-option of hindgut genes in posterior domains at the ectodermal–endodermal interface. The region of the posterior expression of *bra* in the acoel juvenile corresponds to the area at

which the gonoduct will form in the mature adult, which also expresses *bra* (Fig. 2t, u). It is of interest that this position has previously been proposed to represent a common cloacal opening in basal metazoans²⁸.

A molecular dissociation of both mouth and anus formation from the blastopore might explain why the site of mouth formation is distinct from the blastopore in the deuterostomes, and in many lineages of the protostomes¹. Indeed, if the site of gastrulation is defined as the site at which endoderm forms (usually at the vegetal pole), there are few examples in which the mouth forms at this site. The fact that the mouth in acoels does not correspond to the blastopore indicates that the early bilaterian mouth did not form simply by a ventral movement of the blastopore in a planula-like ancestor (Fig. 1b). The position of the mouth, however, varies between acoels from a very anterior position in Hofsteniidae to a posterior site in Diopisthoporidae. Fate map and gene expression studies in these acoels will highlight the developmental changes leading to the variations.

The localization of *gsc* and *bra* expression in the acoel juvenile oral ectoderm also challenges the hypothesis that the protostome and deuterostome larval stages are homologous⁸. Our results indicate that the expression in the foregut of the trochophore and tornaria larva was inherited from a bilaterian ancestor in which a larval form was absent (acoels and nemertodermatids are direct developers). In addition to rejecting hypotheses based on the simultaneous evolution of mouth and anus (for example, the amphistomy-hypothesis²), our results question whether a through gut was present in the last common ancestor of protostomes and deuterostomes because there are

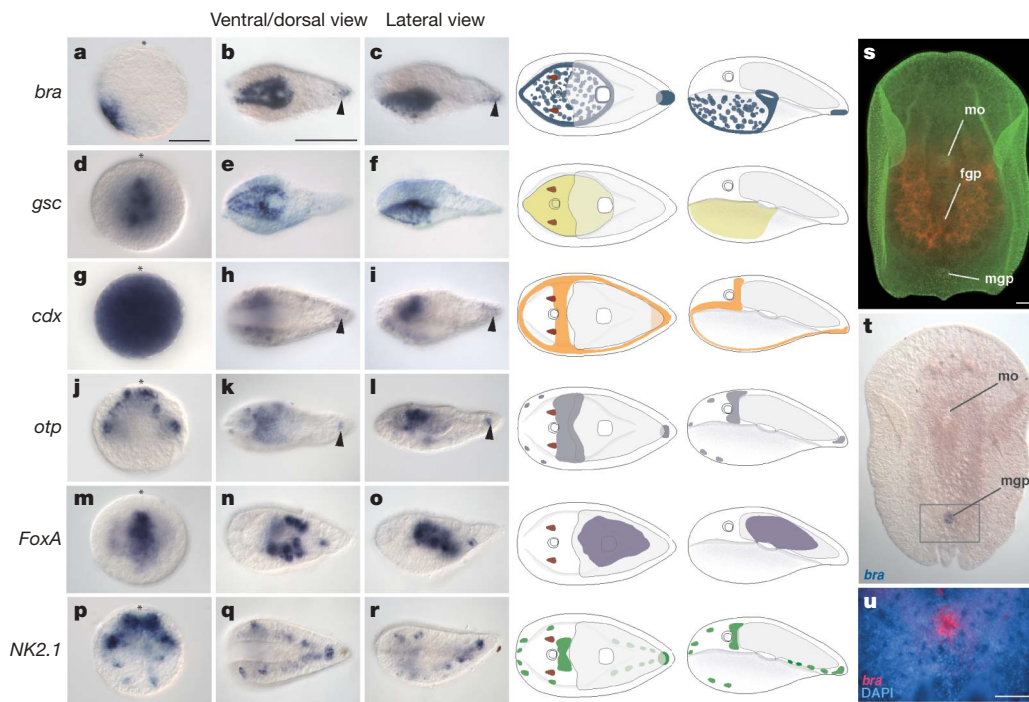


Figure 2 | Gene expression during development of the acoel *C. longifissura*. The first column shows a post-gastrula stage embryo in which the mouth has not yet formed. Asterisks indicate the animal pole, which give rise to the anterior end. The cartoons in the fourth and fifth columns correspond to the juvenile shown in the second and third columns, respectively, and gene expression is coloured. Anterior is to the left, dorsal is to the top. **a**, *bra* is expressed in the future mouth in the post-gastrula; lateral view. **b**, *bra* is expressed in ectoderm around and anterior to the mouth, and in the posterior tip (arrowhead). **c**, Lateral view of **b**. **d**, *gsc* expression in endomesodermal cells. **e**, In the ectoderm, *gsc* is expressed anterior to the mouth opening. **f**, Lateral view of **e**. **g**, Global *cdx* expression in early embryos. **h**, In addition to neural (compare with Fig. 1h) and potential mesodermal tissues, *cdx* is also expressed in the posterior ectoderm

(arrowhead). **i**, Lateral view of **h**. **j**, *otp* is expressed in the post-gastrula in presumptive neural precursor cells. **k**, *otp* is expressed in the juvenile in the brain-commissure and in a posterior spot below the outer ectoderm (arrowhead). **l**, Lateral view of **k**. **m**, *FoxA* is expressed in the endomesoderm of the post-gastrula. **n**, *FoxA* is expressed in the endoderm around the juvenile mouth opening. **o**, Lateral view of **n**. **p–r**, *NK2.1* is expressed in presumptive neural tissue at the anterior end, and in ventral and posterior patches. **s**, Ventral view of an adult stained with Alexa488-labelled phalloidin. Red fluorescence shows the oocytes. Posterior to the mouth (mo) and female gonopore (fgp) is the *bra*-expressing male gonopore (mgp). **t**, In the young adult, *bra* is expressed in the male gonopore. **u**, *bra* gonopore expression of **t**, counterstained with nuclear stain (DAPI); expression shown in red false-colour. Scale bars, 100 μ m.

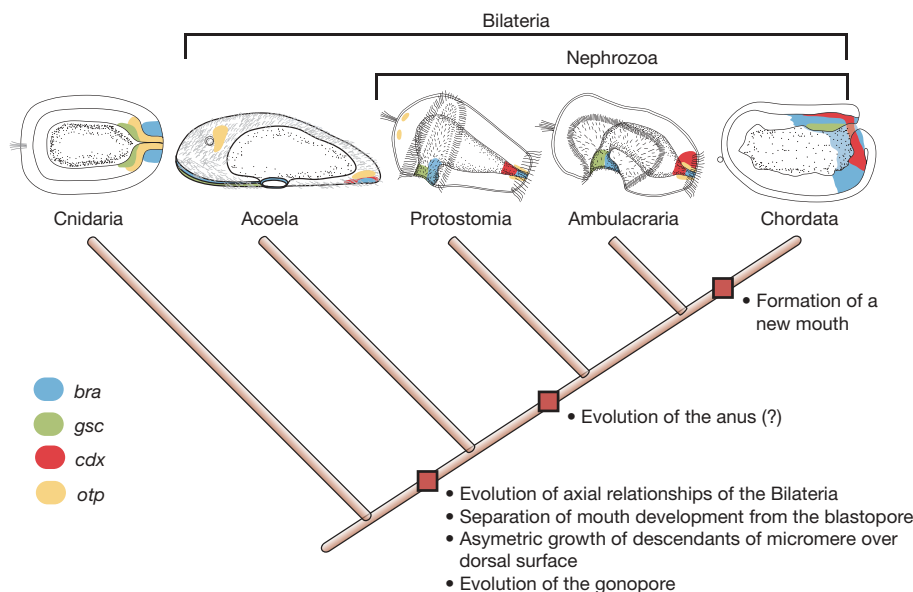


Figure 3 | Position of the Metazoan mouth and anus relative to the expression of *bra*, *gsc*, *cdx* and *otp*. In cnidarians, expression of the genes *bra*, *otp* and *gsc* are located at the 'mouth/blastopore', however, the expression domains are separated in the Bilateria. This separation is independent of the existence of an anus, and the position of the blastopore is as indicated by the data presented for the Acoela. *gsc* and *bra* are

co-expressed at the bilaterian mouth, but *bra*, *cdx* and *otp* are expressed at the posterior end. None of these genes are expressed in the mouth of chordates suggesting the possible independent origin of the mouth cavity in this clade. Only gene expression relevant to the through gut is indicated in this diagram (see Supplementary Table 2 for full description and references). Larval drawings of Protostomia and Ambulacraria are adapted from ref. 8.

basal taxa in both branches that lack an anus (for example, Platyhelminthes and *Xenoturbella*)²⁹. The anal opening of protostomes and deuterostomes could have arisen independently in distinct lineages, which would account for the broader variation of the genes involved in its patterning. The posterior expression domain of *bra* in acoels corresponds to the adult gonopore, which supports suggestions that the anal opening evolved in some animal lineages in association with the somatic reproductive tissue, for example, by forming a cloaca²⁸. However, a further investigation of animals that possess a blind gut, and more detailed analyses of anus formation, will help to explain the course of the evolution of a through gut.

METHODS SUMMARY

Obtaining animals and embryos. Gravid adults were collected from sea tables at the Kewalo Marine Laboratory (Oahu, Hawaii). Embryos were collected at different stages of development up to the juvenile stage.

Gene isolation. Fragments of *bra*, *cdx*, *dlx*, *BMP2/4*, *FoxA*, *gsc*, *NK2.1*, *otp*, *Pitx* and *Six3/6* were amplified using degenerate primers (primer sequences and accession numbers are in Supplementary Table 1). As a template, we used complementary DNA from mixed developmental stages. The amplification of full-length cDNA was carried out using the SMART RACE kit. Gene orthologies were inferred using Bayesian analyses (Supplementary Fig. 4).

Whole mount *in situ* hybridization. Fixation and *in situ* hybridization of embryos and juveniles were performed following an established protocol³⁰, with a shortened proteinase K digestion (2–3 min) and the introduction of a heating step (10 min at 80 °C) before pre-hybridization. Stained embryos and juveniles were imaged with a Zeiss AxioCam HR connected to a Zeiss Axioskop mot2 plus using bright field Nomarski Optics. (Nature Protocols, doi:10.1038/nprot.2008.201.).

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 12 April; accepted 1 August 2008.

Published online 17 September 2008.

- Nielsen, C. *Animal Evolution* 2nd edn (Oxford Univ. Press, 2001).
- Arendt, D. & Nübler-Jung, K. Dorsal or ventral: similarities in fate maps and gastrulation patterns in annelids, arthropods and chordates. *Mech. Dev.* **61**, 7–21 (1997).
- Remane, A. Die Entstehung der Metamerie der Wirbellosen. *Zool. Anz.* **14**, 18–23 (1950).
- Sedgwick, W. On the origin of metamerism and some other morphological questions. *Q. J. Microsc. Sci.* **24**, 43–82 (1884).

- Baguña, J. & Riutort, M. The dawn of bilaterian animals: the case of acoelomorph flatworms. *Bioessays* **26**, 1046–1057 (2004).
- Ruiz-Trillo, I. *et al.* Acoel flatworms: earliest extant bilaterian Metazoans, not members of Platyhelminthes. *Science* **283**, 1919–1923 (1999).
- Wallberg, A., Curini-Galletti, M., Ahmadzadeh, A. & Jondelius, U. Dismissal of Acoelomorpha: Acoela and Nemertodermatida are separate early bilaterian clades. *Zool. Scr.* **36**, 509–523 (2007).
- Arendt, D., Technau, U. & Wittbrodt, J. Evolution of the bilaterian larval foregut. *Nature* **409**, 81–85 (2001).
- Simeone, A. *et al.* *Orthopedia*, a novel homeobox-containing gene expressed in the developing CNS of both mouse and *Drosophila*. *Neuron* **13**, 83–101 (1994).
- Wu, L. H. & Lengyel, J. A. Role of caudal in hindgut specification and gastrulation suggests homology between *Drosophila* amnioproctodeal invagination and vertebrate blastopore. *Development* **125**, 2433–2442 (1998).
- Grobben, K. Die systematische Einteilung des Tierreichs. *Verh. Zool. Bot. Ges. Wien* **58**, 491–511 (1908).
- von Graff, L. *Die Organisation der Turbellaria Acoela* (von Wilhelm Engelmann, 1891).
- Haeckel, E. Die Gastraea-Theorie, die phylogenetische Classification des Tierreiches und die Homologie der Keimblätter. *Jena. Z. Naturwiss.* **8**, 1–55 (1874).
- Hyman, L. H. *The Invertebrates. Vol II. Platyhelminthes and Rhynchocoela* (McGraw-Hill, 1951).
- Jägersten, G. On the early phylogeny of the Metazoa: the bilatero-gastraea theory. *Zool. Bidr. Uppsala* **30**, 321–354 (1955).
- Salvini-Plawen, L. On the origin and evolution of the lower Metazoa. *Z. Zool. Syst. Evolutionsforsch.* **16**, 40–88 (1978).
- Haszprunar, G. Platyhelminthes and Platyhelminthomorpha—paraphyletic taxa. *J. Zool. Syst. Evol. Res.* **34**, 41–48 (1996).
- Marcellini, S. When Brachyury meets Smad1: the evolution of bilateral symmetry during gastrulation. *Bioessays* **28**, 413–420 (2006).
- Tagawa, K., Humphreys, T. & Satoh, N. Novel pattern of *Brachyury* gene expression in hemichordate embryos. *Mech. Dev.* **75**, 139–143 (1998).
- Lartillot, N., Lespinet, O., Vervoort, M. & Adoutte, A. Expression pattern of *Brachyury* in the mollusc *Patella vulgata* suggests a conserved role in the establishment of the AP axis in Bilateria. *Development* **129**, 1411–1421 (2002).
- Neider, A. H., Panopoulou, G. & Langeland, J. A. *Amphioxus gooseoid* and the evolution of the head organizer and prechordal plate. *Evol. Dev.* **2**, 303–310 (2000).
- Fröbisch, A. C. & Seaver, E. C. *ParaHox* gene expression in the polychaete annelid *Capitella* sp. I. *Dev. Genes Evol.* **216**, 81–88 (2006).
- Lowe, C. J. *et al.* Anteroposterior patterning in hemichordates and the origins of the chordate nervous system. *Cell* **113**, 853–865 (2003).
- Christiansen, L. *et al.* Evolutionary modification of mouth position in deuterostomes. *Semin. Cell Dev. Biol.* **18**, 502–511 (2007).
- Henry, J. Q., Martindale, M. Q. & Boyer, B. C. The unique developmental program of the acoel flatworm, *Neochilidia fusca*. *Dev. Biol.* **220**, 285–295 (2000).

26. Denes, A. S. *et al.* Molecular architecture of annelid nerve cord supports common origin of nervous system centralization in bilateria. *Cell* **129**, 277–288 (2007).
27. Lowe, C. J. *et al.* Dorsoventral patterning in hemichordates: insights into early chordate evolution. *PLoS Biol.* **4**, e291 (2006).
28. Reisinger, E. Allgemeine Morphologie der Metazoa. *Fortschr. Zool.* **13**, 1–82 (1961).
29. Dunn, C. W. *et al.* Broad phylogenomic sampling improves resolution of the animal tree of life. *Nature* **452**, 745–749 (2008).
30. Martindale, M. Q., Pang, K. & Finnerty, J. R. Investigating the origins of triploblasty: 'mesodermal' gene expression in a diploblastic animal, the sea anemone *Nematostella vectensis* (phylum, Cnidaria; class, Anthozoa). *Development* **131**, 2463–2474 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work is supported by the National Science Foundation under the ATOL program (EF05-31558) and NASA to M.Q.M. A.H. received further support from the Deutsche Forschungsgemeinschaft (HE5183/2-1). We thank K. Pang for improving the manuscript and A. Okuso for the drawings of the cnidarian and annelid.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to A.H. (hejmol@hawaii.edu).

METHODS

Scanning electron microscopy. To prepare the juveniles for scanning electron microscopy, the animals were relaxed in 7% magnesium chloride and fixed in ice cold 4% paraformaldehyde in sea water for 5 min, followed by fixation in 2.5% glutaraldehyde in 0.1 M sodium cacodylate buffer with 0.35 M sucrose for 1 h. After post-fixation in 1% osmium tetroxide, the animals were dehydrated and critical point dried, sputtered and imaged with a Hitachi S-800 scanning electron microscope.

Confocal laser scanning microscopy. After relaxation in magnesium chloride, juveniles were fixed in ice cold 4% paraformaldehyde in sea water for 1 h and washed five times with PBS containing 0.1% Triton X-100 and incubated with an anti-tyrosinated-tubulin antibody (Sigma, dilution 1:500) overnight. After five washes in PBS, the juveniles were incubated in the Cy3-labelled secondary antibody against mouse immunoglobulin-M (Jackson Laboratories). Animals were stained with Alexa488-labelled phalloidin (Molecular Probes) for 1 h and washed three times in PBS. After staining with 4,6-diamidino-2-phenylindole (DAPI), the animals were mounted in glycerol and imaged with a Zeiss LSM 510.

Four-dimensional microscopy. Freshly deposited egg clusters were mounted in sea water on a microscope slide, with the coverslip sealed with vaseline to prevent evaporation. Using a Zeiss Imager Z1 microscope and a camera (Hamamatsu Orca ER) triggered by Volocity software (Improvision), images were taken every minute. Images were exported and combined to generate a time-lapse movie using Final Cut Pro (Apple).

Sequencing the nuclear genome of the extinct woolly mammoth

Webb Miller¹, Daniela I. Drautz¹, Aakrosh Ratan¹, Barbara Pusey¹, Ji Qi¹, Arthur M. Lesk¹, Lynn P. Tomsho¹, Michael D. Packard¹, Fangqing Zhao¹, Andrei Sher^{2,†}, Alexei Tikhonov³, Brian Raney⁴, Nick Patterson⁵, Kerstin Lindblad-Toh⁵, Eric S. Lander⁵, James R. Knight⁶, Gerard P. Irzyk⁶, Karin M. Fredrikson⁷, Timothy T. Harkins⁷, Sharon Sheridan⁷, Tom Pringle⁸ & Stephan C. Schuster¹

In 1994, two independent groups extracted DNA from several Pleistocene epoch mammoths and noted differences among individual specimens^{1,2}. Subsequently, DNA sequences have been published for a number of extinct species. However, such ancient DNA is often fragmented and damaged³, and studies to date have typically focused on short mitochondrial sequences, never yielding more than a fraction of a per cent of any nuclear genome. Here we describe 4.17 billion bases (Gb) of sequence from several mammoth specimens, 3.3 billion (80%) of which are from the woolly mammoth (*Mammuthus primigenius*) genome and thus comprise an extensive set of genome-wide sequence from an extinct species. Our data support earlier reports⁴ that elephantid genomes exceed 4 Gb. The estimated divergence rate between mammoth and African elephant is half of that between human and chimpanzee. The observed number of nucleotide differences between two particular mammoths was approximately one-eighth of that between one of them and the African elephant, corresponding to a separation between the mammoths of 1.5–2.0 Myr. The estimated probability that orthologous elephant and mammoth amino acids differ is 0.002, corresponding to about one residue per protein. Differences were discovered between mammoth and African elephant in amino-acid positions that are otherwise invariant over several billion years of combined mammalian evolution. This study shows that nuclear genome sequencing of extinct species can reveal population differences not evident from the fossil record, and perhaps even discover genetic factors that affect extinction.

Vertebrate genome sequencing projects have thus far assembled data from at least 28 species⁵, including chromosomal assemblies of six placental mammals, namely human^{6,7}, chimpanzee⁸, rhesus macaque⁹, mouse¹⁰, rat¹¹ and dog¹². In contrast, kilobase (kb)-scale genomic sequence data from extinct species were first reported in 2005, with 27 kb from cave bear¹³ and 13 megabases (Mb) from mammoth¹⁴. More recently, two projects reported up to 1 Mb from Neanderthal^{15,16}, some of which may be modern human contamination¹⁷.

Whereas many ancient-DNA studies have used bone samples, in 2007 we showed that DNA with fewer damage-induced substitutions can be extracted from hair shafts collected from permafrost remains¹⁸. Moreover, use of hair permits a highly efficient decontamination protocol that leaves the keratin-encased endogenous DNA unharmed. The method resulted in 15 complete mammoth mitochondrial genomes at high coverage^{18,19}, identified in 947 Mb of total

sequence (average read length, 93 bp) from a number of samples. Hair shafts are thus suitable for sequencing ancient nuclear DNA.

We selected M4, a Siberian mammoth specimen carbon-14 dated to $18,545 \pm 70$ years before present (roughly 20,000 years ago), for extensive sequencing, and generated 2.982 Gb of data from hair shafts using a Roche GS-FLX sequencing instrument. A second mammoth specimen, M25, yielded an additional 239 Mb. Together with our earlier mammoth data, this brought the total to 4.168 Gb of sequence. Given the abundance of hair available from M4 and M25, we were able to enrich the sequenced material for long DNA molecules, to overcome at least partly the high rate of breakage in ancient DNA. The average read length was 142 bp for M4 and 164 bp for M25. As a bonus, we obtained 4,430-fold coverage of the mitochondrial genome of M4, allowing us to determine error rates. (We assumed that errors in nuclear DNA equal those in mitochondrial DNA.) Specifically, for reads trimmed by aligning them to elephant sequence, the total error rate from post-mortem DNA damage and sequencing mistakes was 0.14%, neglecting errors that added or deleted bases (Table 1 and Methods).

To estimate how many of our reads are actual mammoth DNA, we determined the fraction of our sequence that aligns to the African savanna elephant (*Loxodonta africana*) genome (twofold assembly and sixfold reads), which indicated that 80% of our 4.17 Gb of sequence, that is, approximately 3.3 Gb, is from the mammoth. However, the yield varied substantially between specimens: M4 is 90% mammoth and M25 is only 58% mammoth (Fig. 1). As a negative control, read-sized intervals of the chicken genome²⁰ were mapped to elephant and showed that the false-positive rate is very low. Reasons why the estimation of 80% may actually be low are given in Methods. Some microbial DNA is recognizable in the non-mammoth portion (Fig. 1), but essentially none of the DNA in these samples is human¹⁸.

The converse result, that is, the fraction of elephant DNA that aligns to our data, can tell us how much of the mammoth genome has been sequenced. Because typical genome sizes of placental mammals are around 3 Gb, our 3.3 Gb might be expected to provide at least onefold coverage, in which case—taking into account overlapping reads²¹—over 63% of the bases in the mammoth genome would be sequenced at least once. However, the African elephant genome has previously been estimated at between 4.2 and 4.8 Gb using the C-value technique⁴, which, although less accurate than genome sequencing, has consistently predicted the Afrotherian genomes to be larger than previously sequenced placental genomes. We estimated how much of

¹Pennsylvania State University, Center for Comparative Genomics and Bioinformatics, 310 Wartik Building, University Park, Pennsylvania 16802, USA. ²Severtsov Institute of Ecology and Evolution, Russian Academy of Sciences, 33 Leninsky Prospekt, 119071 Moscow, Russia. ³Zoological Institute, Russian Academy of Sciences, Universitetskaya Naberezhnaya 1, 199034 Saint Petersburg, Russia. ⁴Center for Biomolecular Science and Engineering, University of California, Santa Cruz, California 95064, USA. ⁵Broad Institute of MIT and Harvard, Cambridge Center, Cambridge, Massachusetts 02142, USA. ⁶454 Life Sciences, 20 Commercial Street, Branford, Connecticut 06405, USA. ⁷Roche Diagnostics Corporation, 9115 Hague Road, Indianapolis, Indiana 46250-0414, USA. ⁸Sperling Foundation, Eugene, Oregon 97405, USA.

†Deceased

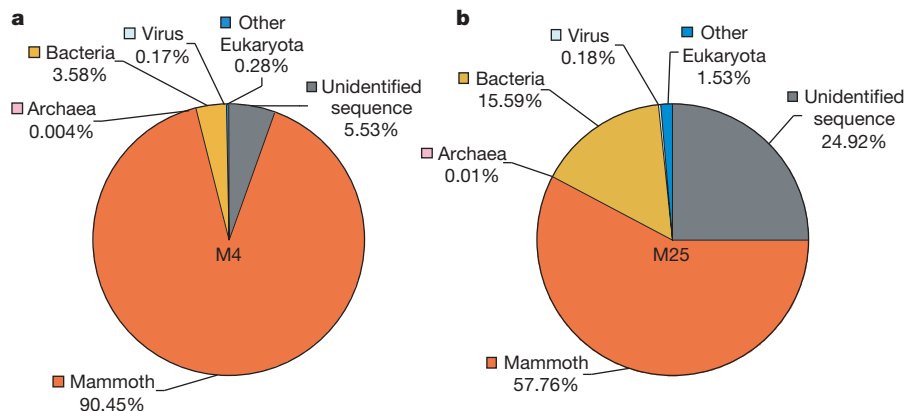


Figure 1 | Species composition of metagenomic DNA extracted from mammoth hair. a, b, Pie charts for the M4 (a) and M25 (b) data sets show the percentage of sequencing reads assigned to taxa for mammoth, Archaea,

Bacteria, virus, as well as the two unspecified categories ‘other Eukaryota’ and ‘unidentified sequence’. The taxon distribution exemplifies the variability of the endogenous DNA content of ancient specimens.

the mammoth genome has been sequenced by searching for matches to a set of elephant genes in the Ensembl gene build of August 2006 (<http://www.ensembl.org>) that could be confidently mapped to unique positions on human chromosomes (Fig. 2), and by searching for the so-called ultraconserved regions²². In both cases, around 50% of the bases were found; accounting for multiple reads that include the same genomic position, this translates into 0.7-fold coverage, or that the total length of our true mammoth reads is 70% of the genome's length. Because some of our reads are very short and, hence, difficult to align reliably, this may be an underestimate.

Our estimates that 80% of our 4.17 Gb of sequence is from mammoth and that we have obtained 0.7-fold coverage are consistent with a genome size of 4.7 Gb, as $4.17 \times 0.8 \approx 3.3 \approx 4.7 \times 0.7$. However, this estimation of genome size is at best a rough approximation. On the other hand, we observed the probable cause of the expanded genome, namely an unusually high fraction of interspersed repeats (Supplementary Information).

As currently understood, the evolutionary relationships among selected living and extinct elephantid species are sketched in Fig. 3; we show parallels with humans and some great apes to provide a widely familiar point of reference. Here we use estimated divergence times, which are times to the common ancestor averaged across the genome. This should be distinguished from population split times or, in the case of distinct species, speciation times. For instance, two modern European humans have a population split time of 0 yr but a mean divergence of at least 500,000 yr. This distinction is important for recent speciation events. For example, the mean divergence time between human and chimpanzee is at least 2 Myr longer than the speciation time²³ (see Methods for details). We are interested in comparing sequence identity rates between elephantids and between apes.

Table 1 | Basic statistics on the mammoth genome sequence

Sequenced bases	4.168 Gb
Sequencer runs (Roche GS-FLX and GS20)	77
Sequenced reads	32,619,456
Average read length	128 bp
Bases that align to <i>L. africana</i>	3.3 Gb
Sequence coverage for M4's mitochondrial genome	4,430-fold
Total error rate based on mitochondrial genome	35 per 10,000 bp
Substitutions from DNA damage	6 per 10,000 bp
Substitutions from sequencing error	8 per 10,000 bp
Insertions/deletions from sequencing error	21 per 10,000 bp
Total error neglecting indels	14 per 10,000 bp (0.14%)
Estimated genome size	4.7 Gb
Estimated nucleotide identity of M4 to African elephant	99.41%
Estimated amino-acid identity of M4 to African elephant	99.78%

To estimate the level of nucleotide identity (ignoring gaps) between M4 and the African elephant, we analysed the large number of elephant positions that have more than one aligning mammoth read, to reduce the effect of errors in our sequence (Methods). The estimated identity is 99.4%. The 0.6% difference rate is approximately half of that estimated between human and chimpanzee (1.24%)²⁴, despite the similarity in divergence times (Fig. 3). This indicates that nucleotide substitutions are fixed in recent elephantid lineages at only half of the rate in great apes and humans, mirroring an earlier observation about mitochondrial DNA²⁵. Using a similar approach (Methods), we estimate that M4 and the African elephant are 99.78% identical at the amino-acid level.

Significantly, among the multiply sequenced differences between M4 and the African elephant, M25 agrees with the African elephant in 13.3% of the cases (that is, 327 of 2,451) in which we had a high-identity alignment to M25. Under the assumption that M4 and the African elephant differ by 15 Myr of evolution (7.5 Myr in each lineage), this corresponds to a separation of about 1.5–2.0 Myr between

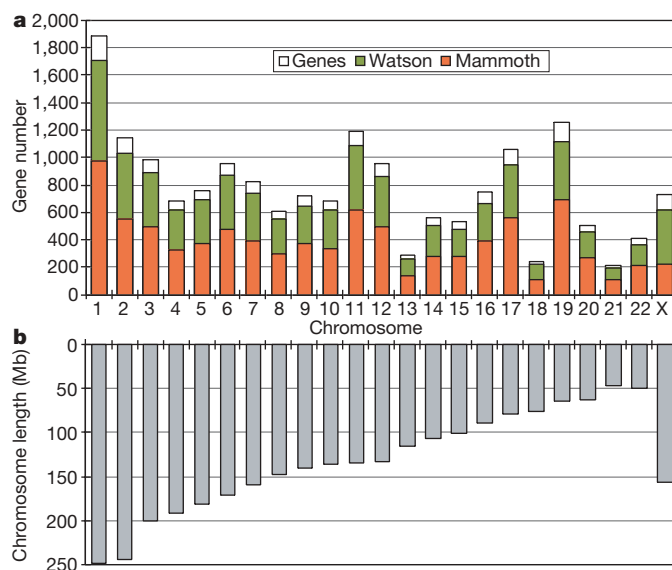


Figure 2 | Sequenced mammoth orthologues of human genes. a, Plot showing the number of RefSeq genes on each human chromosome (open white rectangles), the average fraction of protein-coding bases that align to Roche/454 reads from James D. Watson's genome³⁰ (green), and the fraction of coding bases that align to one or more mammoth reads (red), using Ensembl-predicted elephant genes that map to the human chromosome—approximately 50% for each autosome, but only 31% for chromosome X as M4 was male (see Methods). b, Lengths of the chromosomes in a.

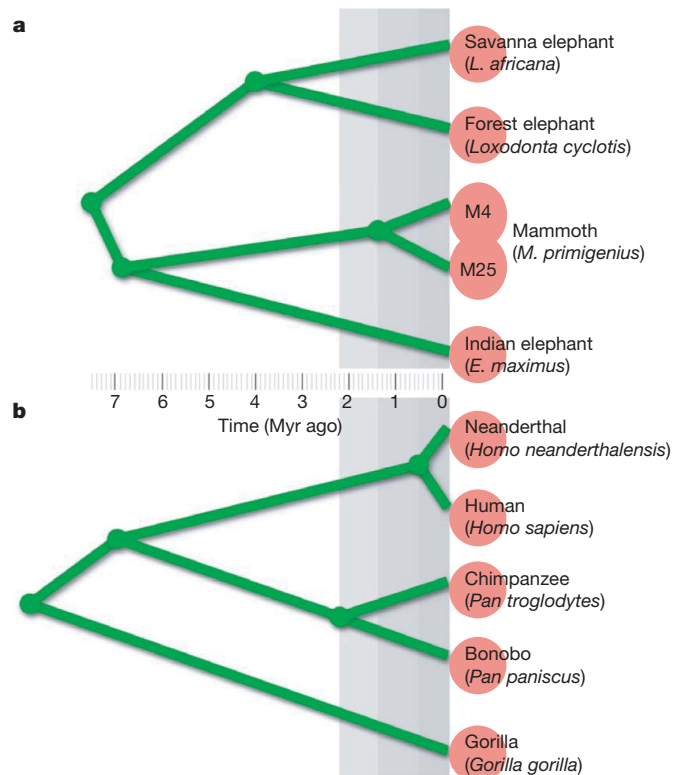


Figure 3 | Comparison of phylogenies. **a**, Elephantids; **b**, hominoids. We show estimated divergence times, that is, times to the common ancestor averaged across autosomes (see Methods). Red circles at the leaves of the phylogenetic tree indicate discernable species. This distinction was not made for the two clades of mammoth (M4 and M25) based on the fossil record (merged red circles).

M4 and M25. We assume that only a small fraction of the differing positions are under selection. We note that a divergence of 1–2 Myr between M4 and M25 was estimated earlier¹⁹ on the basis of mitochondrial data, for which M25 agrees with the African elephant in 14.5% of the cases (114 of 792) where M4 disagrees with the African elephant. The concordance between nuclear and mitochondrial data

is particularly noteworthy because population-genetic analysis of African elephants has shown that different relationships are inferred from mitochondrial sequence than from nuclear sequence²⁶. M4 and M25 belong to differing clades of mammoths that were identified on the basis of short mitochondrial sequences^{19,27}. However, morphological criteria distinguishing the two clades have not been established, similar to the case of two phenotypically identical groups of extant brown bears in Sweden that have differing mitotypes and share the same territory²⁸.

A major reason for sequencing the woolly mammoth is to identify functionally important amino-acid differences between mammoth and elephant. It is unclear what fraction of amino-acid differences have functional consequences, but it is likely to be rather small; for instance, one estimation²⁹ is that ~20% of common human amino-acid variants are deleterious. To start looking for such differences, we combined computational criteria (designed to enrich for validity and functional importance) with PCR amplification and Sanger sequencing in M4, M25, African elephant and Indian elephant, *Elephas maximus*. Our initial screening yielded 92 putative differences (Fig. 4, Supplementary Information) that have also passed additional manual screening for undesirable attributes such as lack of conservation (notably homoplasy) at the critical mammoth position, potential confusion with paralogues, processed and unprocessed pseudogenes, and tandem or other duplicative debris. We found a number of cases in which mammoth differed from an amino acid that appeared to be otherwise invariant throughout placental evolution (Supplementary Information), which may suggest functional significance of the protein position and positive selection in the mammoth lineage.

From the data set presented here, we conclude that a high-fidelity, high-coverage mammoth genome will be feasible once the genome sequence for the African elephant has been completed and 10–30-fold (depending on the sequencing technology) more mammoth sequence has been generated. From our data, we estimate that mammoth and elephant differ on average at about one residue per protein (roughly 20,000 positions proteome-wide) and that 90% of those differences are potentially identifiable by means of higher-coverage short-read sequencing alone (that is, without enriching sequenced material for coding DNA or Sanger resequencing; see Methods). Apart from comparing protein sequences, we hope to pinpoint DNA differences between mammoth and elephant in the non-repetitive genomic

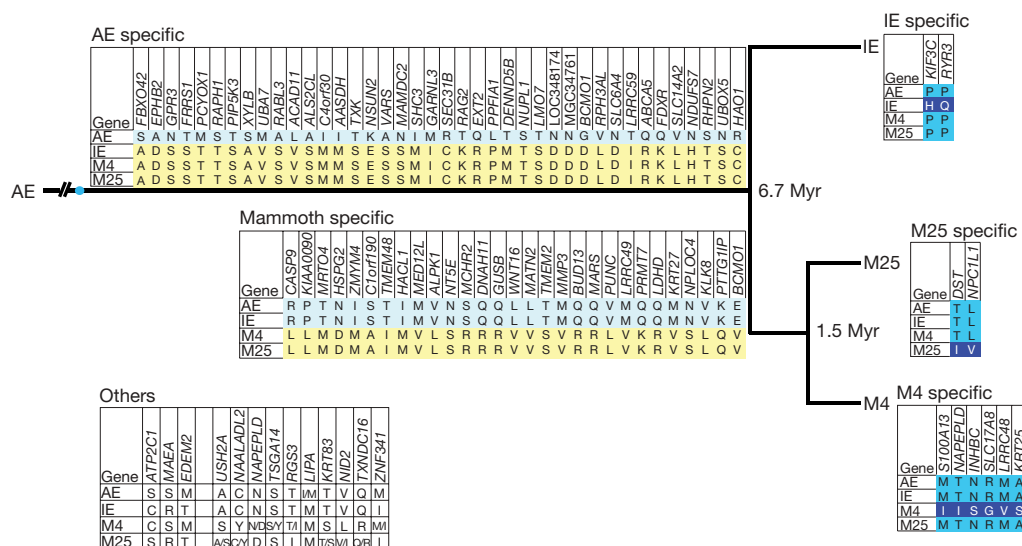


Figure 4 | Experimentally verified amino-acid differences among African elephant, Indian elephant, M4 and M25. Non-synonymous coding single-nucleotide polymorphisms between African elephant and mammoth identified by computational mean, termed single-amino-acid polymorphisms (SAPs), were experimentally verified through PCR

amplification and Sanger sequencing. Six SAP categories for splits specific to African elephant (AE), Indian elephant (IE), mammoth, M4 and M25 are shown, together with one category for heterozygosity and other polymorphisms. Gene names are for the putative human orthologue.

intervals, so it may even be possible to detect differences in gene-regulatory signals. The catalogue of differences, along with computational predictions of the differences most likely to have functional consequences, will provide a resource to facilitate direct observation of genetically orchestrated changes over evolutionary time, for example those associated with adaptation to cold environments, dietary changes and so on. In addition, the determination of an even larger number of synonymous changes in those protein-coding intervals will permit identification of genes and gene families under selection during mammoth evolution. As we have demonstrated here, these studies can be carried out on both clades of mammoth despite the specimens' large difference in age. It will therefore become possible to conduct genome-wide studies on multiple individuals with the goal of understanding whether the observed coalescence time of 1.5–2.0 Myr between M4 and M25 in fact did generate two species of mammoth, or whether this process was terminated by premature extinction of the clade of M25.

METHODS SUMMARY

The 4.17 Gb of individual reads from our mammoth samples, along with the sequenced PCR products produced while studying SAPs, were placed in a freely Internet-accessible BLAST server (<http://mammoth.psu.edu>) that was used for some of the analyses described here, such as determining the fraction of putative elephant coding exons and mammalian ultraconserved intervals that align to a mammoth-sample read. In addition, sequence data from the African savanna elephant genome, produced by the Broad Institute, was a critical ingredient for our analysis. Early in the project, we used an assembly of the twofold-coverage data, downloaded from the University of California, Santa Cruz Genome Bioinformatics website (<http://genome.ucsc.edu>); later we employed individual Sanger reads that provide roughly sixfold coverage.

For whole-genome mammoth–elephant comparisons, we handled the problem of unmasked elephantid-specific interspersed repeats by aligning mammoth-sample reads to the twofold elephant scaffolds using the 'dynamic masking' feature of the BLASTZ alignment program (see Methods); only the reads that did not align in that preliminary step were aligned to the sixfold reads. Reads that aligned to a unique position in the twofold assembly, and specifically to a position itself aligned with high identity to a human RefSeq coding exon, were analysed to predict elephant/mammoth SAPs. We assumed that any given read is either all mammoth DNA or completely non-mammoth. One set of computational conditions, designed to enrich for substitutions of functional importance, was (1) that the putative amino-acid difference between mammoth and elephant occur in a run of five amino acids identical between human and elephant, and (2) that the substitution have a negative BLOSUM62 score.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 12 May; accepted 22 September 2008.

1. Höss, M., Pääbo, S. & Vereshchagin, N. K. Mammoth DNA sequences. *Nature* **370**, 333 (1994).
2. Hagelberg, E. *et al.* DNA from ancient mammoth bones. *Nature* **370**, 333–334 (1994).
3. Pääbo, S. *et al.* Genetic analyses from ancient DNA. *Annu. Rev. Genet.* **38**, 645–679 (2004).
4. Redi, C. A. *et al.* Genome size: a novel genomic signature in support of Afrotheria. *J. Mol. Evol.* **64**, 484–487 (2007).
5. Miller, W. *et al.* 28-way vertebrate alignment and conservation track at the UCSC Genome Browser. *Genome Res.* **17**, 1797–1808 (2007).
6. International Human Genome Sequencing Consortium. Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).
7. Venter, J. C. *et al.* The sequence of the human genome. *Science* **291**, 1304–1351 (2001).

8. The Chimpanzee Sequencing and Analysis Consortium. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437**, 69–87 (2005).
9. Rhesus Macaque Genome Sequencing and Analysis Consortium. Evolutionary and biomedical insights from the rhesus macaque genome. *Science* **316**, 222–234 (2007).
10. Mouse Genome Sequencing Consortium. Initial sequencing and comparative analysis of the mouse genome. *Nature* **420**, 520–562 (2002).
11. Rat Genome Sequencing Project Consortium. Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature* **428**, 493–541 (2004).
12. Lindblad-Toh, K. *et al.* Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature* **438**, 803–819 (2005).
13. Noonan, J. P. *et al.* Genomic sequencing of Pleistocene cave bears. *Science* **309**, 597–600 (2005).
14. Poinar, H. N. *et al.* Metagenomics to palaeogenomics: large-scale sequencing of mammoth DNA. *Science* **311**, 392–394 (2006).
15. Noonan, J. P. *et al.* Sequencing and analysis of Neanderthal genomic DNA. *Science* **314**, 1113–1118 (2006).
16. Green, R. E. *et al.* Analysis of one million base pairs of Neanderthal DNA. *Nature* **444**, 330–336 (2006).
17. Wall, J. D. & Kim, S. K. Inconsistencies in Neanderthal genomic DNA sequences. *PLoS Genet.* **3**, e175 (2007).
18. Gilbert, M. T. P. *et al.* Whole-genome shotgun sequencing of mitochondria from ancient hair shafts. *Science* **317**, 1927–1930 (2007).
19. Gilbert, M. T. P. *et al.* Intraspecific phylogenetic analysis of Siberian woolly mammoths using complete mitochondrial genomes. *Proc. Natl Acad. Sci. USA* **105**, 8327–8332 (2008).
20. International Chicken Genome Sequencing Consortium. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* **432**, 695–716 (2004).
21. Lander, E. S. & Waterman, M. S. Genomic mapping by fingerprinting random clones: A mathematical analysis. *Genomics* **2**, 231–239 (1988).
22. Bejerano, G. *et al.* Ultraconserved elements in the human genome. *Science* **304**, 1321–1325 (2004).
23. Patterson, N. *et al.* Genetic evidence for complex speciation of humans and chimpanzees. *Nature* **441**, 1103–1108 (2006).
24. Chen, F. C. & Li, W. H. Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. *Am. J. Hum. Genet.* **68**, 444–456 (2001).
25. Rohland, N. *et al.* Proboscidean mitogenomics: chronology and mode of elephant evolution using mastodon as an outgroup. *PLoS Biol.* **5**, e207 (2007).
26. Roca, A. L. *et al.* Cytonuclear genomic dissociation in African elephant species. *Nature Genet.* **37**, 96–100 (2005).
27. Barnes, I. *et al.* Genetic structure and extinction of the woolly mammoth, *Mammuthus primigenius*. *Curr. Biol.* **17**, 1072–1075 (2007).
28. Taberlet, P. *et al.* Localization of a contact zone between two highly divergent mitochondrial DNA lineages of the brown bear *Ursus arctus* in Scandinavia. *Conserv. Biol.* **9**, 1255–1261 (1995).
29. Sunyaev, S. *et al.* Prediction of deleterious human alleles. *Hum. Mol. Genet.* **10**, 591–597 (2001).
30. Wheeler, D. A. *et al.* The complete genome of an individual by massively parallel DNA sequencing. *Nature* **452**, 872–876 (2008).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This sequencing-by-synthesis study was made possible through generous funding from Penn State University and Roche Applied Sciences. W.M. was supported by grant HG002238 from the National Human Genome Research Institute and S.C.S. is supported in part by the Gordon and Betty Moore Foundation. This project is funded in part under a grant from the Pennsylvania Department of Health using Tobacco Settlement Funds appropriated by the US legislature. The Pennsylvania Department of Health specifically disclaims responsibility for any analyses, interpretations or conclusions. We thank T. Gilbert for introducing us to the use of hair shafts as a source of ancient DNA, C. Grøndahl for providing Asian elephant hair samples, and M. Wilson for suggestions. This paper is dedicated to the memory of Andrei Sher.

Author Information Sequence data are available at the NCBI database's trace archive section under accession number SRA001810. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to W.M. (webb@bx.psu.edu) or S.C.S. (scs@bx.psu.edu).

METHODS

Sample preparation. We employed the protocols described in a previous paper¹⁸, augmented as follows. M4 and M25 DNA were size-selected before DNA library construction by running the samples on a 2% unstained agarose gel along with a 100-bp DNA ladder (N3231S, New England Biolabs). The ladder was excised and stained for fragment visualization and the M4 and M25 DNA were excised between the 400-bp and 1,000-bp fragments that corresponded to the 100-bp DNA ladder. The samples were then purified using a QIAquick Gel Extraction Kit (QIAGEN) and used for library construction according to the manufacturer's instructions (Roche Applied Sciences).

Error rate. We used the 4,430-fold coverage of the M4 mitochondrial genome (see Supplementary Fig. 1) to assess the sequencing error and post-mortem DNA damage to our sample, by comparing the individual reads with their consensus. We observed 72,951,869 matches, 103,181 mismatches, 51,587 erroneous missing bases and 97,661 erroneous extra bases (Supplementary Table 1). The total error rate was 0.345%. The most frequent damage-induced substitutions^{31,32} are C → T and G → A, with G → A most probably an artefact of the T4 polymerase enzyme in the DNA library construction³³. Following an earlier paper³¹ we estimated the C → T DNA damage rate by subtracting the number of T → C transitions from the number of C → T transitions, and similarly for G → A. Together these account for 17.5% of the total errors, whereas other (putatively sequencing error) mismatches constitute 23.4%. The remaining 59.1% of the error consists of insertions/deletions in the reads, mostly in homopolymer runs. Thus, ignoring erroneous insertions and deletions, the fraction of incorrect bases is predicted to be 40.9% (that is, 17.5% + 23.4%) of 0.345%, or about 0.14%.

The fraction of reads from the mammoth genome. We randomly picked a set equivalent to a half-run on the Roche/454 instrument (roughly 30 Mb). The reads were aligned to the twofold elephant assembly using the BLASTZ program³⁴ (scoring matches, 1; mismatches, -3; unaligned nucleotides, -1; local alignment of score, ≥30). Once a read was found to align to the elephant assembly, it was not compared to subsequent elephant sequences, to avoid aligning each interspersed or tandem repeat segment thousands of times. Reads that did not align to the twofold elephant assembly were compared with the sixfold elephant (Sanger) reads, requiring a gap-free alignment scoring of at least 30 when a match of 1 and a mismatch of -3. The reads that aligned to *L. africana* in one of these two steps, comprising 80.1% of the 4.17 Gb, were considered to be mammoth DNA. We also used this approach on just the M4 sample, just the M25 sample, and on read-sized intervals from the chicken genome²⁰; the latter indicated that 1.2% of non-mammoth reads and 1.6% of their bases align to elephant under our criteria.

On the other hand, there are at least two reasons why the value of 80% may underestimate the amount of mammoth DNA. First, as about 1% of the human genome sequence is contained in a read-sized interval that is completely absent from chimpanzee, and vice versa⁸, 1% of the mammoth genome may be missed by our method because there is no orthologous *L. africana* sequence. Also, the Roche/454 technology that we used can sequence parts of the genome that are not represented in Sanger-based data such as that for elephant, for example an additional 1.4% of the human genome³⁰.

Non-mammoth reads. The metagenomic results summarized in Fig. 1 (see Supplementary Information for details) were obtained using the MEGAN³⁵ metagenome analysis software (<http://www-ab.informatik.uni-tuebingen.de/software/megan/welcome.html>).

The fraction of the genome that is sequenced. We downloaded 15,717 Ensembl elephant gene models (ftp://ensembl.org/pub/release-49/fasta/loxodonta_africana/cdna) and determined a subset of 4,131 that could be computationally assigned to a unique orthologous position in the human genome. These were compared with mammoth-sample reads using the mammoth BLAST server (<http://mammoth.psu.edu>), and alignments were filtered to require at least 97% identity. For autosomal genes, 51% of the elephant protein-coding bases were aligned to one or more mammoth reads by the nucleotide-based aligner blastn, where we required alignments to have at least 97% identity. On the other hand, only 30.9% of the coding bases from the X chromosome aligned (Fig. 2), reflecting the fact that M4 was male (see below); it is very likely that some of the other sequences were from females, which explains why the fraction for the X chromosome is a little more than half the autosomal fraction.

In another test, we aligned mammoth reads to the 481 so-called ultraconserved regions²²—481 intervals (mostly not protein-coding) averaging 262 bp that are identical among human, mouse and rat and, hence, very likely to be observable in any other placental mammal. These intervals total 121,250 bp, and we aligned them using BLASTZ with the parameters $T = 2$, $Y = 2,000$, a scoring match of 100 and a mismatch of -200, and penalized a gap of length N by $400 + 100N$. This time, we identified 47% of the bases in our data.

M4 was male. We identified a number of M4 reads that align to male-specific human genes³⁶. For instance, we found a 286-bp read that aligns at 84% identity

to the *HSFY1* mRNA (GenBank accession no. NM_033108). As anticipated, neither the read nor the *HSFY1* mRNA aligns to the current elephant assembly (sequenced from a female), even under much less stringent conditions.

Divergence times. For Fig. 3, the estimated divergence times within elephantids are²⁵ as follows: African elephants diverged from mammoths and Asian elephants approximately 7.6 Myr ago, mammoths diverged from Indian elephants approximately 6.7 Myr ago, and African savannah elephants diverged from African forest elephants (*Loxodonta cyclotis*) approximately 4.0 Myr ago. The divergence time of the two mammoth clades, represented by M4 and M25, is estimated to be 1.5 Myr, on the basis of ref. 19 and data presented in this paper. For the human–Neanderthal divergence, we use a date of 400,000 years, concordant with the fossil record, and add 500,000 years for within-Neanderthal divergence, similar to that for modern humans. For the other hominoid divergence times, we use^{23,37} 7.0 Myr for human and chimpanzee, 2.2 Myr for chimpanzee and bonobo (*Pan paniscus*) and 8.75 Myr for human and gorilla.

The nucleotide similarity between M4 and elephant. We limited consideration to M4 reads that aligned to a single position in the elephant assembly, and further required that the aligning portion be at least 100 bp and align with at least 97% identity. We looked for elephant positions (1) that aligned to more than one mammoth read, (2) that always aligned to the same nucleotide and (3) where neither the elephant nucleotide nor an aligning nucleotide is ambiguous (that is, 'N'). (This approach discarded elephant positions that aligned to non-identical mammoth nucleotides; such positions accounted for 0.28% of the multiply aligned elephant positions, consistent with the prediction that if two mammoth nucleotides are aligned to a given elephant position, then one of them will be incorrect because of sequencing error or DNA damage (and hence the two will differ) about 0.28% of the time, as each one has a predicted substitution-error rate of 0.14%.) Finally, we required that the elephant position be aligned to mammoth reads that were generated on different days, to avoid the possibility of artefactual duplicates caused by our sample-preparation protocol. We found 45,039,470 such consistently aligning positions, of which 44,773,945 (99.41%) were identical between mammoth and elephant. To be more precise, this is the estimated rate of identity between mammoth and the current elephant assembly. Among the consistently aligning positions of M4 and African elephant, 2,451 were covered by M25 reads that aligned exactly once to elephant, with at least 97% identity over at least 100 bp; M5 agreed with elephant in 327 cases.

The amino-acid similarity between M4 and elephant. We identified 175,949 pairwise non-overlapping human RefSeq protein-coding exons (9,911,624 amino acids, not counting those split by an intron), where we required that the exon has the same reading frame in all annotated splice variants that contain it. Of these exons, 99,946 (4,284,551 amino acids) align without a gap or an improper stop codon to the twofold elephant assembly available in May 2007, using human–elephant alignments downloaded from the University of California, Santa Cruz Genome Bioinformatics browser³⁸. Of these exons, 76,750 (containing 3,331,646 amino acids) had at least 85% amino-acid identity between human and elephant. To help eliminate matches between paralogues (but retaining orthologues), we worked with that reduced set.

Orthologous mammoth amino acids were predicted as follows. We limited consideration to 454-sequencer reads that aligned to a unique position of the elephant assembly, where we required at least 97% nucleotide identity (counting gaps) between elephant and the aligned portion of the read. Whenever a read overlapped an exon and, within the overlap, there were no gaps or 'N's in elephant or mammoth, we recorded the human-, elephant- and mammoth-aligned amino acids in the overlap (nucleotides before or after the reading frame were ignored). This procedure produced 1,537,885 amino-acid triples, including cases where a human–elephant position aligns to multiple mammoth reads. As with nucleotide differences (see above), we identified putative elephant amino-acid positions that appeared more than once among these triples, and were always aligned to the same predicted mammoth amino acid. There were 165,532 such cases, of which 165,164 (99.78%) were identical between mammoth and elephant.

Amino-acid differences. The Supplementary Information contains (1) our computational protocol for picking potentially interesting amino-acid differences between African elephant and M4, (2) our experimental methods for validating the predictions, (3) detailed information about the validated difference and (4) data about evolutionary conservation of the differences highlighted in Supplementary Table 5.

How many mammoth–elephant amino-acid differences can potentially be determined? Human RefSeq genes contain just under 10 million amino acids (removing redundancy caused by splice isoforms), and hence we estimate the corresponding number for elephant to be 10 million. Assuming 99.8% amino-acid identity between elephant and mammoth, this gives roughly 20,000 amino-acid differences. To detect a mammoth–elephant difference using short reads (say onefold coverage with 200-bp reads and sixfold coverage with 40-bp reads),

it is necessary that the read can be mapped to the elephant assembly with high confidence. This becomes difficult in portions of the assembly that align elsewhere in the assembly, say with at least 95% identity over at least 100 bp. We determined that roughly 9% of the human RefSeq coding intervals meet these conditions. On this basis, we anticipate that roughly 90% of our reads that intersect a protein-coding region can be reliably mapped to the completed elephant assembly.

31. Stiller, M. *et al.* Patterns of nucleotide misincorporations during enzymatic amplification and direct large-scale sequencing of ancient DNA. *Proc. Natl Acad. Sci. USA* **103**, 13578–13584 (2006).
32. Gilbert, M. T. P. *et al.* Recharacterization of ancient DNA miscoding lesions: Insights in the era of sequencing-by-synthesis. *Nucleic Acids Res.* **35**, 1–10 (2007).
33. Brotherton, P. *et al.* Novel high-resolution characterization of ancient DNA reveals C > U-type base modification events as the sole cause of post mortem miscoding lesions. *Nucleic Acids Res.* **35**, 5717–5728 (2007).
34. Schwartz, S. *et al.* Human-mouse alignments with Blastz. *Genome Res.* **13**, 103–107 (2003).
35. Huson, D. H. *et al.* MEGAN analysis of metagenomic data. *Genome Res.* **17**, 377–386 (2007).
36. Skaletsky, H. *et al.* The male-specific region of the human Y chromosome is a mosaic of discrete sequence classes. *Nature* **423**, 825–837 (2003).
37. Caswell, J. L. *et al.* Analysis of chimpanzee history based on genome sequence alignments. *PLoS Genet.* **4**, e1000057 (2008).
38. Kent, W. J. *et al.* The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).

Guarding the gateway to cortex with attention in visual thalamus

Kerry McAlonan¹, James Cavanaugh¹ & Robert H. Wurtz¹

The massive visual input from the eye to the brain requires selective processing of some visual information at the expense of other information, a process referred to as visual attention. Increases in the responses of visual neurons with attention have been extensively studied along the visual processing streams in monkey cerebral cortex, from primary visual areas to parietal and frontal cortex^{1–4}. Here we show, by recording neurons in attending macaque monkeys (*Macaca mulatta*), that attention modulates visual signals before they even reach cortex by increasing responses of both magnocellular and parvocellular neurons in the first relay between retina and cortex, the lateral geniculate nucleus (LGN). At the same time, attention decreases neuronal responses in the adjacent thalamic reticular nucleus (TRN). Crick⁵ argued for such modulation of the LGN by observing that it is inhibited by the TRN, and suggested that “if the thalamus is the gateway to the cortex, the reticular complex might be described as the guardian of the gateway”, a reciprocal relationship we now show to be more than just hypothesis. The reciprocal modulation in LGN and TRN appears only during the initial visual response, but the modulation of LGN reappears later in the response, suggesting separate early and late sources of attentional modulation in LGN.

We recorded responses of LGN and TRN neurons in three awake behaving macaque monkeys. Monkeys were directed by a central cue at the point of fixation to attend to one of two peripheral visual stimuli on randomly interleaved trials (Fig. 1a inset). One of these stimuli was in the receptive field (RF) of the recorded neuron. Figure 1a shows the responses of an example magnocellular LGN neuron (LGNm) to a light bar within the RF when attention was directed out of the RF (dashed curve, ATTout) or into the RF (solid curve, ATTin). Responses shown are from correct trials. The ATTin response falls above the ATTout response, indicating an increase in neuronal response with attention. The mean response to the same stimulus increased 12% with attention. Figure 1b shows the responses of a parvocellular LGN neuron (LGNp) that also increased (21%) when attention was directed into the RF.

If a similar increase in attention were to occur in TRN, however, Crick's hypothesized interaction between TRN and LGN encounters a problem: TRN inhibits LGN. The visual sector of TRN receives excitatory inputs from LGN, but projects modulatory inhibitory input back to LGN^{6–13}. Therefore, TRN responses should instead decrease with attention, reducing the inhibitory influence of the TRN on LGN, thereby causing the increase in the responses of LGN neurons that we observe. We did in fact find a decrease in the TRN visual response with attention (Fig. 1c). When attention was directed into the RF of this TRN neuron, the mean response to the same visual stimulus was 13% less than when attention was directed out of the RF.

We have summarized the effect of attention on mean visual responses of 57 on-centre LGN neurons (19 LGNm and 38 LGNp) in

Fig. 2a, and of 29 TRN neurons in Fig. 2b. In each plot, the ordinate is the baseline ATTout response and the abscissa is the attentional modulation, ATTmod. ATTmod can be expressed either as the contrast measure $(ATTin - ATTout)/(ATTin + ATTout)$, or the ratio of modulation $(ATTin/ATTout)$. We have included both, with the

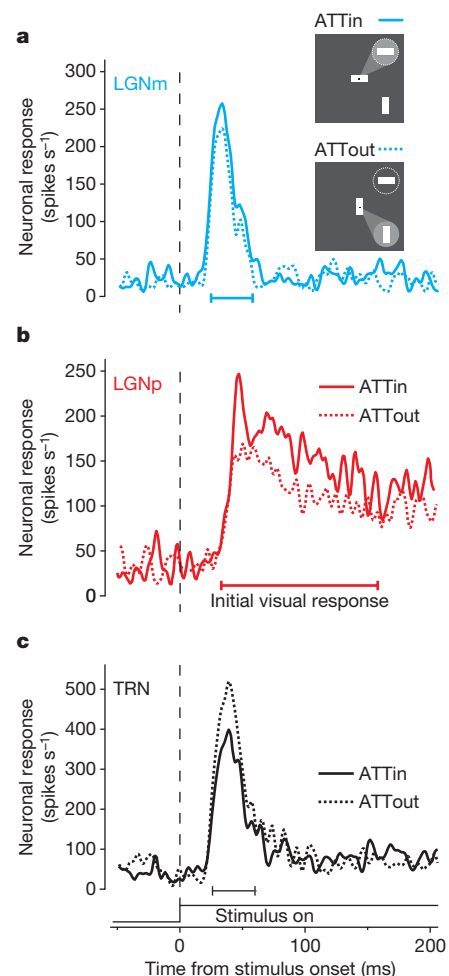


Figure 1 | Sample responses to shifts of attention in LGN and TRN. Solid traces are spike density plots of the neuron's ATTin response (as illustrated by the 'spotlight' of attention in the inset cartoon directed to the circle representing the RF). Dashed traces are ATTout responses. Responses are aligned to stimulus onset (the dashed vertical line), and have been smoothed with a Gaussian window of 2 ms s.d. **a**, Responses of sample magnocellular LGN neuron (LGNm). **b**, Responses of sample parvocellular LGN neuron (LGNp). **c**, Responses of sample TRN neuron.

¹Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20892, USA.

bottom axes representing the ratio of modulation, and the top axes indicating the contrast measure of ATTmod. Figure 2a shows the bulk of points to the right of the vertical unity line, indicating that the predominant effect of attention in LGN neurons was to increase mean responses to the visual stimulus. Distributions of ATTmod appear above the scatter plot, with small arrows indicating sample medians. In LGN, attention increased the median response $11 \pm 2.6\%$ in the magnocellular layers ($P = 0.011$), and $9 \pm 1.1\%$ in the parvocellular layers ($P = 0.0007$). All indications of variability are ± 1 standard error of the median, and all P values were determined using the Wilcoxon signed rank test for zero median, unless otherwise specified.

In contrast to LGN, values of attentional modulation in TRN (Fig. 2b) tend to lie to the left of the unity line, showing a median decrease in neuronal response with attention of $4 \pm 0.6\%$ ($P = 0.004$). Over our sample of neurons, the reciprocal effect of attention holds; LGN responses increase with attention whereas TRN responses decrease.

If attention modulates neuronal responses, we would not necessarily expect such modulation during trials on which the monkeys made incorrect behavioural responses. For TRN neurons with more than five error trials, responses on those trials increased by $1.5 \pm 1.5\%$ ($n = 18$, $P = 0.62$). Similarly, for LGNm and LGNp neurons, responses changed by $2.3 \pm 3.6\%$ ($n = 11$, $P = 0.58$) and $1.3 \pm 2.6\%$ ($n = 22$, $P = 0.71$). The lack of significant

response modulation on incorrect trials provides further evidence that the factor enabling the monkeys to perform the task correctly was the same one modulating neuronal responses, namely, visual attention.

We found no significant modulation of the background activity preceding the initial visual response or in the latency or duration of the initial visual response in either LGN or TRN. To observe any residual effect of attention beyond the initial visual response, we examined mean neuronal responses from 100 ms before stimuli appeared to 500 ms after they appeared (the shortest presentation time common to all trials). For each neuron, we normalized the response to the neuron's maximum firing rate. Figure 3a shows the mean normalized response for each area with solid curves for the ATTin condition and dashed curves for the ATTout condition. We calculated ATTmod for the six 100-ms time epochs in this timescale. Figure 3b shows ATTmod (as ATTin/ATTout) over time. Median changes for each area are connected across epochs with solid lines, and error bars denote ± 1 standard error of the median. Significant changes within an epoch are denoted by coloured asterisks.

All areas demonstrate a significant response modulation in the initial 100 ms epoch after the stimuli appear. However, this modulation disappears in the next 100 ms epoch, but LGNm and LGNp show a second, later period of modulation that becomes significant in both divisions as time progresses. Also, both LGNm and LGNp showed significant attentional modulation just before the monkey needed to make a decision about the stimulus. Note that, in contrast to LGN, TRN had no second period of attentional modulation.

Because only the initial visual response in TRN is modulated by attention, measuring over the whole 500 ms period would have yielded a much smaller modulation in TRN (-1.8%) that would not have been significant ($P = 0.31$). However, owing to the second phase of modulation in LGN, we still would have measured attentional modulation of 13% in LGNm ($P = 0.014$) and 8.1% in LGNp ($P < 0.0001$), but the influence of TRN on the initial visual response would have gone undetected.

We see that careful consideration of responses over time is critical to detect the attentional effects in TRN, but analysis of the interactions between LGN and TRN requires an even more precise examination of

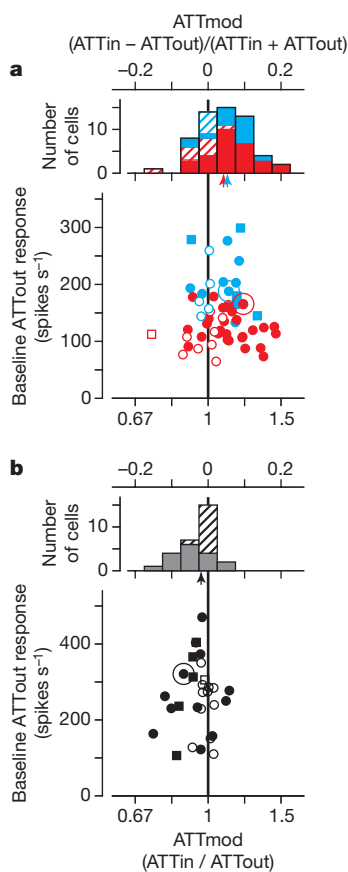


Figure 2 | Effect of attention on LGN and TRN. **a**, Scatter plot showing mean baseline ATTout response versus attentional modulation (ATTmod) for 19 LGNm neurons (blue) and 38 LGNp neurons (red). Solid symbols are significant response changes (Wilcoxon rank-sum test, $P < 0.01$). Squares denote experiments in which performance did not guarantee attention (see Methods). Distributions of ATTmod appear above the scatter plots. Hatched areas of each bar denote changes that did not reach significance. Arrows show median ATTmod for LGNm (blue) and LGNp (red). **b**, Similar scatter plot and histogram for TRN. In all plots, the larger circles indicate the Fig. 1 example neurons.

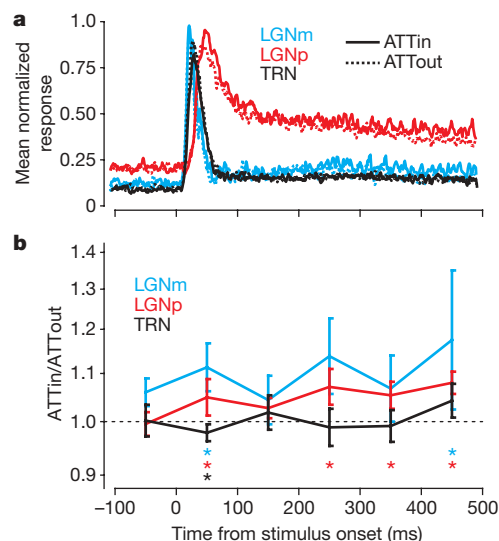


Figure 3 | Time courses of visual and attentional influences. **a**, Mean normalized ATTin (solid curves) and ATTout (dashed curves) responses for each area. Each curve is the mean normalized spike density plot over all neurons in an area. Mean responses have been smoothed with a Gaussian kernel of 2.8 ms s.d. **b**, Median effect of attention on each area in 100 ms epochs. Each trace shows ATTmod over time. Error bars are ± 1 standard error of the median. Significant changes in each epoch are denoted by coloured asterisks coded to each area below the curves.

response timing. To compare visual response latencies, Fig. 4a shows the mean normalized initial responses for neurons in each area aligned on stimulus onset. To determine the significance of visual latency differences, we performed a bootstrap analysis (see Supplementary Notes) yielding estimates of the median visual latency in each area, and the significance of differences between areas using the Wilcoxon rank-sum test for equal medians. Although TRN responses (median latency 22 ± 0.92 ms) begin well before those in LGNp ($P < 0.0001$, median latency 37 ± 1.43 ms), LGNm neurons (median latency 21 ± 1.25 ms) tend to respond before TRN neurons ($P < 0.0001$).

To track the timing of attentional modulation in each area, we represented the effect of attention in Fig. 4b as the difference between the mean ATTin and ATTout curves from Fig. 4a. The latency of attentional modulation was obtained from a similar bootstrap analysis. Whereas the visual response appears first in LGNm, Fig. 4b shows that attentional modulation occurs first in TRN (22 ± 0.37 ms), 4 ms before LGNm (26 ± 0.31 ms, $P < 0.0001$). The attentional effect shows up significantly later in LGNp (37 ± 0.31 ms) than either TRN or LGNm ($P < 0.0001$ for both). Therefore, even though LGNm visual responses precede those of TRN (consistent with LGNm driving the visual response in TRN), attention affects TRN responses first, consistent with attentional modulation in LGN coming from TRN.

In summary, we find that attention modulates thalamic visual responses in two phases: an initial modulation that attenuates TRN responses and enhances LGN responses, followed by a slowly building later enhancement limited to LGN. Until now, demonstration of attentional modulation of LGN neurons has been limited to preliminary experiments on monkey¹⁴ and fMRI studies in humans¹⁵. For the TRN, in addition to the recent growth in anatomical and cellular studies of monkey visual TRN^{6,8,9,13}, we recently found attentional modulation of neuronal activity in visual TRN during a visual/auditory attention task¹⁶. The differences between the visual/auditory attention task and the current task, along with a comparison of their results, are given in the Supplementary Discussion.

The initial LGN modulation might provide a substantial fraction of the modulation seen subsequently in cortical area V1 (refs 17–23). Although it is difficult to compare across studies, the approximately 10% increase in responses we find in LGN is similar to the 6.9% median increase across V1 neurons¹⁷, and the 8.9% median increase in V1

simple cells¹⁸. The presence of the initial modulation in both TRN and LGN, their reciprocal increase and decrease, and the timing of their visual and attentional responses are consistent with TRN serving as the source of the initial LGN modulation, as proposed by Crick.

The later attentional effects in LGN, and effects others have reported in higher cortical visual areas, might be more closely related to goal-directed attention, which frequently also develops later in the visual response particularly in higher cortical areas^{2,4,24}. This later modulation in LGN might in fact reflect feedback from cortex onto the LGN^{25,26} via the established connections from V1 layer 6 (refs 25, 27), whereas the initial modulation in LGN by way of TRN may have its origins in subcortical structures, possibly including the superior colliculus^{28–30}. Although obviously separate in time course, the two phases of modulation may represent two distinct attentional influences, and may be early indicators for identifying and distinguishing feed-forward and feedback visual attentional mechanisms.

METHODS SUMMARY

Two monkeys performed a task in which a central cue directed them to attend to one of two peripheral stimuli (a horizontal and a vertical bar of light; Fig. 1a inset). On each trial, while the monkey fixated on a central spot, a cue appeared at the fixation point matching one of the two upcoming stimuli. On any trial, the cue had an equal chance of matching either the vertical or horizontal stimulus. After 250 ms, the two peripheral stimuli appeared, one in the RF of the neuron, and the other some distance from the RF. After a period of 500–1,000 ms, each peripheral stimulus independently had a 50% chance of transiently dimming about 40% in luminance. The monkey indicated if the stimulus matching the cue dimmed by making a saccade to it. If the matching stimulus did not dim, the correct response was to remain fixating. The correct response depended only on the stimulus matching the central cue. We compared neuronal responses when the cue matched the stimulus in the RF (ATTin) with responses when the cue matched the remote stimulus (ATTout). For each neuron studied, we always presented the same stimulus in the RF, and only the cue changed randomly between trials. This ensured that neurons responded to the same stimulus regardless of which stimulus matched the central cue. Eye movements were monitored to make certain the monkey remained fixating during trials. The possible contribution of changes in the monkeys' eye position on the attentional modulation is considered in Supplementary Discussion. Because the stimulus in the RF always remained the same, this also allowed the monkey to shift attention as soon as the cue came on, as the location of each stimulus was consistent from trial to trial.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 18 July; accepted 27 August 2008.

Published online 5 October 2008.

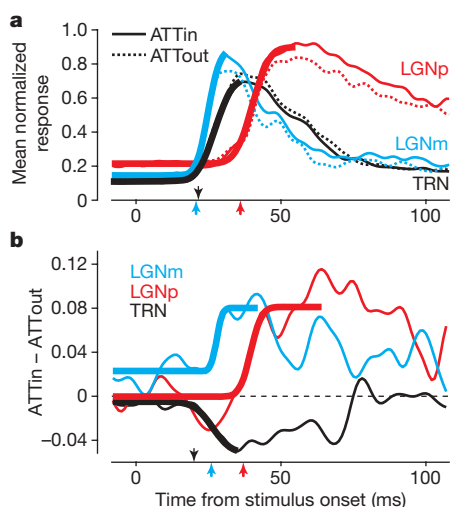


Figure 4 | Latencies of visual and attentional influences. **a**, First 100 ms of the visual responses from Fig. 3a. Thick lines are sample descriptive fits to ATTin responses from the bootstrap analysis. Arrows indicate median visual latencies. **b**, Latency of attentional effect in each area. Each trace shows the difference between the mean ATTin and ATTout responses. Thick lines show sample descriptive fits used to extract latency estimates in the bootstrap analysis. Arrows indicate the median latencies for the effect of attention.

- Desimone, R. & Duncan, J. Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**, 193–222 (1995).
- Maunsell, J. H. & Cook, E. P. The role of attention in visual processing. *Phil. Trans. R. Soc. Lond. B* **357**, 1063–1072 (2002).
- McAdams, C. J. & Maunsell, J. H. Effects of attention on the reliability of individual neurons in monkey visual cortex. *Neuron* **23**, 765–773 (1999).
- Reynolds, J. H. & Chelazzi, L. Attentional modulation of visual processing. *Annu. Rev. Neurosci.* **27**, 611–647 (2004).
- Crick, F. Function of the thalamic reticular complex: The searchlight hypothesis. *Proc. Natl Acad. Sci. USA* **81**, 4586–4590 (1984).
- Conley, M. & Diamond, I. T. Organization of the visual sector of the thalamic reticular nucleus in *Galago*. *Eur. J. Neurosci.* **2**, 211–226 (1990).
- Crabtree, J. W. & Killackey, H. P. The topographic organization and axis of projection within the visual sector of the rabbit's thalamic reticular nucleus. *Eur. J. Neurosci.* **1**, 94–109 (1989).
- Fitzgibbon, T., Szmajda, B. A. & Martin, P. R. First order connections of the visual sector of the thalamic reticular nucleus in marmoset monkeys (*Callithrix jacchus*). *Vis. Neurosci.* **24**, 857–874 (2007).
- Harting, J. K., Van Lieshout, D. P. & Feig, S. Connectional studies of the primate lateral geniculate nucleus: Distribution of axons arising from the thalamic reticular nucleus of *Galago crassicaudatus*. *J. Comp. Neurol.* **310**, 411–427 (1991).
- Jones, E. G. Some aspects of the organization of the thalamic reticular complex. *J. Comp. Neurol.* **162**, 285–308 (1975).
- Kim, U., Sanchez-Vives, M. V. & McCormick, D. A. Functional dynamics of GABAergic inhibition in the thalamus. *Science* **278**, 130–134 (1997).
- Montero, V. M., Guillery, R. W. & Woolsey, C. N. Retinotopic organization within the thalamic reticular nucleus demonstrated by a double label autoradiographic technique. *Brain Res.* **138**, 407–421 (1977).

13. Uhlrich, D. J., Manning, K. A. & Feig, S. L. Laminar and cellular targets of individual thalamic reticular nucleus axons in the lateral geniculate nucleus in the prosimian primate *Galago*. *J. Comp. Neurol.* **458**, 128–143 (2003).
14. Casagrande, V. A., Sary, G., Royal, D. & Ruiz, O. On the impact of attention and motor planning on the lateral geniculate nucleus. *Prog. Brain Res.* **149**, 11–29 (2005).
15. O'Connor, D. H., Fukui, M. M., Pinsk, M. A. & Kastner, S. Attention modulates responses in the human lateral geniculate nucleus. *Nature Neurosci.* **5**, 1203–1209 (2002).
16. McAlonan, K., Cavanaugh, J. & Wurtz, R. H. Attentional modulation of thalamic reticular neurons. *J. Neurosci.* **26**, 4444–4450 (2006).
17. McAdams, C. J. & Maunsell, J. H. Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J. Neurosci.* **19**, 431–441 (1999).
18. McAdams, C. J. & Reid, R. C. Attention modulates the responses of simple cells in monkey primary visual cortex. *J. Neurosci.* **25**, 11023–11033 (2005).
19. Motter, B. C. Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *J. Neurophysiol.* **70**, 909–919 (1993).
20. Roelfsema, P. R., Lamme, V. A. & Spekreijse, H. Object-based attention in the primary visual cortex of the macaque monkey. *Nature* **395**, 376–381 (1998).
21. Ito, M. & Gilbert, C. D. Attention modulates contextual influences in the primary visual cortex of alert monkeys. *Neuron* **22**, 593–604 (1999).
22. Posner, M. I. & Gilbert, C. D. Attention and primary visual cortex. *Proc. Natl Acad. Sci. USA* **96**, 2585–2587 (1999).
23. Roberts, M., Delicato, L. S., Herrero, J., Gieselmann, M. A. & Thiele, A. Attention alters spatial integration in macaque V1 in an eccentricity-dependent manner. *Nature Neurosci.* **10**, 1483–1491 (2007).
24. Treue, S. Neural correlates of attention in primate visual cortex. *Trends Neurosci.* **24**, 295–300 (2001).
25. Briggs, F. & Usrey, W. M. A fast, reciprocal pathway between the lateral geniculate nucleus and visual cortex in the macaque monkey. *J. Neurosci.* **27**, 5431–5436 (2007).
26. Sillito, A. M., Cudeiro, J. & Jones, H. E. Always returning: feedback and sensory processing in visual cortex and thalamus. *Trends Neurosci.* **29**, 307–316 (2006).
27. Lund, J. S., Lund, R. D., Hendrickson, A. E., Bunt, A. H. & Fuchs, A. F. The origin of efferent pathways from the primary visual cortex, area 17, of the macaque monkey as shown by retrograde transport of horseradish peroxidase. *J. Comp. Neurol.* **164**, 287–303 (1975).
28. Jones, E. G. in *The Thalamus* 1270–1271 (Cambridge Univ. Press, 2007).
29. Kolmac, C. I. & Mitrofanis, J. Patterns of brainstem projection to the thalamic reticular nucleus. *J. Comp. Neurol.* **396**, 531–543 (1998).
30. Wilson, J. R., Hendrickson, A. E., Sherk, H. & Tigges, J. Sources of subcortical afferents to the macaque's dorsal lateral geniculate nucleus. *Anat. Rec.* **242**, 566–574 (1995).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was supported by the intramural research program of the National Eye Institute. We are grateful for the assistance of J. McClurkin, A. Nichols, M. Smith, T. Ruffner and G. Tansey.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to K.M. (km@nei.nih.gov).

METHODS

Physiological methods. One recording chamber allowed access to both TRN and LGN. It was implanted stereotaxically 10 mm anterior from the interaural line and 13 mm lateral from the midline on three male rhesus macaque monkeys (monkey B, O and G—monkey G provided only TRN data). The surgical procedures, recording of single neurons and eye positions, and control of the monkeys' behaviour have been described previously¹⁶. All procedures were approved by the Institute Animal Care and Use Committee and complied with Public Health Service Policy on the humane care and use of laboratory animals.

We initially localized LGN using an MRI image after the recording cylinder was implanted. LGN recordings were verified by the nature of the visual response and the signature alternation of the ocularity of these responses as the electrode progressed through layers. TRN was located by its position in relation to LGN. The short latency of the visual response (about 22 ms) and its brief duration (~60 ms) identified the neurons as being from TRN as determined previously with histological verification¹⁶, rather than from the parvocellular layers in dorsal LGN.

For most units, RF centre and extent were quantitatively determined. While the monkey fixated on a central spot 0.4° in diameter, a spot of light (0.8° to 1.0°) appeared sequentially in a series of locations centred on the estimated RF centre, arranged in a 5×5 grid. Horizontal and vertical separation between grid points was equal to the diameter of the spot. Each trial lasted as long as 800 ms and the spot appeared in each randomly chosen location for either 200 ms (LGN) or 100 ms (TRN). Responses to the spot at each location were analysed online and the centre of the RF was adjusted accordingly. We subsequently determined RF diameter in a similar manner, this time by presenting a sequence of spots with varying diameters centred on the RF. Spots were 0.5° to 6° in diameter, and appeared in random sequence during each 800 ms trial, each spot remaining for 200 ms (LGN) or 100 ms (TRN). The quantitatively determined RF centre and diameter were then used to place the stimuli for the attention task. Visual stimuli in all tasks were back projected onto a tangent screen 58 cm in front of the monkey by a liquid crystal display projector.

Attention task details. During a trial, the monkey was required to maintain fixation within 1.0° or 1.5° of the central fixation point (Supplementary Fig. 1). If the monkey broke fixation early, the trial was aborted and excluded from analysis. The central cue was $1.5^\circ \times 0.6^\circ$ or $2.0^\circ \times 0.8^\circ$, whichever was closest to the size of the peripheral stimuli. The peripheral stimuli were 1.5° to 2.0° long and 0.6° to 0.8° wide, depending on the eccentricity and size of the RF. One of the stimuli was placed in the RF of the LGN or TRN neuron. The other stimulus was

placed some distance away, typically about 20° , but at the same eccentricity from the fixation point as the stimulus in the RF. Each peripheral bar of light independently had a 50% chance of dimming for 600 ms. Therefore on 25% of trials both stimuli dimmed (simultaneously), on 25% of trials neither stimulus dimmed, on 25% of trials only the horizontal stimulus dimmed, and on 25% of trials only the vertical stimulus dimmed.

We used a criterion of 75% correct responses to indicate that the monkey was attending to the cued target, although performance was typically better. To flag possible response strategies, we first divided the trials into eight trial types according to which stimulus was cued (horizontal or vertical) and which stimuli dimmed (horizontal only, vertical only, both, or neither). If the monkey followed some response strategy rather than attending as directed by the cue, performance on one or more of these trial types would necessarily suffer. We flagged possible response strategies by requiring the monkey get each of these eight trial types correct at least 50% of the time. So when we refer to the monkey performing the attention task to criteria, the monkey is getting at least 75% of the trials correct overall and is getting at least 50% of each type of trial correct.

Analysis of results. We measured the activity of neurons as the mean neuronal response within several different epochs. We measured mean background activity in the 100 ms period before onset of the visual stimuli. Visual response latency was determined by fitting a normal cumulative density function (CDF) to the spike density plot obtained over at least 20 but usually more than 50 trials and smoothed with a 2.8 ms s.d. Gaussian kernel (Supplementary Fig. 2a). The onset of the visual response was taken as the time at which the fit curve reached 10% of the neuron's peak response. The latency of this response was the time between onset of the visual stimulus (determined by a photo cell attached to the screen) and this response onset time. The end of the initial response was taken as the point at which a Weibull probability density function (PDF) fit to the falling phase of the response declined 75% from the neuron's peak response to the asymptote of the fit curve. The duration of the initial visual response was then the time between the visual response onset and the end of the initial visual response. Both the gaussian CDF and the Weibull PDF were fitted by minimizing the sum-squared error between the fit curve and the spike density plots. The mean response is the average response rate during this epoch.

Comparisons with other studies. The percentage response modulation we report for other studies was calculated from available firing rates with baseline activity included. From the neuronal response rates with and without attention, we were able to extract a comparable percentage change from the ratio of modulation by calculating $100 \times [(ATT_{in}/ATT_{out}) - 1]$.

Mechanism of phototaxis in marine zooplankton

Gáspár Jékely¹†, Julien Colombelli², Harald Hausen³, Keren Guy¹, Ernst Stelzer², François Nédélec² & Detlev Arendt¹

The simplest animal eyes are eyespots composed of two cells only: a photoreceptor and a shading pigment cell. They resemble Darwin's 'proto-eyes', considered to be the first eyes to appear in animal evolution^{1–4}. Eyespots cannot form images but enable the animal to sense the direction of light. They are characteristic for the zooplankton larvae of marine invertebrates and are thought to mediate larval swimming towards the light. Phototaxis of invertebrate larvae contributes to the vertical migration of marine plankton⁵, which is thought to represent the biggest biomass transport on Earth^{6,7}. Yet, despite its ecological and evolutionary importance, the mechanism by which eyespots regulate phototaxis is poorly understood. Here we show how simple eyespots in marine zooplankton mediate phototactic swimming, using the marine annelid *Platynereis dumerilii* as a model⁸. We find that the selective illumination of one eyespot changes the beating of adjacent cilia by direct cholinergic innervation resulting in locally reduced water flow. Computer simulations of larval swimming show that these local effects are sufficient to direct the helical swimming trajectories towards the light. The computer model also shows that axial rotation of the larval body is essential for phototaxis and that helical swimming increases the precision of navigation. These results provide, to our knowledge, the first mechanistic understanding of phototaxis in a marine zooplankton larva and show how simple eyespots regulate it. We propose that the underlying direct coupling of light sensing and ciliary locomotor control was a principal feature of the proto-eye and an important landmark in the evolution of animal eyes.

We assayed phototaxis of *Platynereis* larvae (Fig. 1a, b) in a dark chamber with >750 nm transmission illumination. Larval swimming in a 1×1×0.2 cm cuvette in the dark chamber was recorded after providing a directional diffuse white light stimulus at one end of the chamber (Supplementary Fig. 1 and Supplementary Movie 1). We measured the efficiency of phototaxis by calculating the average displacement of the larvae towards the light (Fig. 1c). Positive phototaxis started at 25 h of development and lasted for two days. The spectral sensitivity was broad (365–545 nm) with two maxima at 410 and 490 nm (Fig. 1d), similar to that of phototactic larvae from other species⁹. Phototaxis could be triggered with intensities >0.5 mW cm⁻² using 500-nm monochromatic light. Phototactic larvae swim in a right-handed helix (that is, clockwise forward movement) of a diameter of 60–300 µm while rotating around their anterior–posterior axis with one full turn every 0.6–0.9 s. During swimming the dorsal side of the larva is always directed towards the axis of the helix so that a 360° axial rotation takes place in one pitch of the helix (Fig. 1e). This is a general phenomenon among protozoa and ciliated animal larvae^{10–14}.

Platynereis larvae develop a pair of eyespots consisting of a pigment cell and a rhabdomeric photoreceptor cell^{8,15} (Fig. 1b, f). The pigment cup of the pigment cell is directly adjacent to the photosensitive rhabdom of the photoreceptor and shades it from one side (Fig. 1f–i). Both eyespots have wide, slightly dorsally and posteriorly tilted planes of

view as determined by electron microscopy three-dimensional reconstructions of individual pigment vesicles and of the rhabdom outline (Fig. 1h–j). The restricted view angle of the photoreceptors confers differential sensitivity to the direction of light that is thought to be the basis of phototaxis^{12,16}. To test whether it is the eyespots and not the developing adult eyes or brain ciliary photoreceptors^{8,17} that mediate phototaxis, we performed bilateral ablations of eyespots using pulsed laser nanosurgery^{18,19} (Supplementary Movie 2). This resulted in the complete loss of phototaxis (0 phototactic larvae out of 40 ablated ones) as opposed to ablation of two cells in the apical organ (27 phototactic larvae out of 32 ablated ones). In contrast, when we ablated only one eyespot most larvae were still able to swim towards the light (left eye ablated: 19 out of 24, right eye ablated: 13 out of 21 phototactic). This shows that eyespots work as independent units and that the comparison of light levels between the eyespots is not needed during phototaxis.

Platynereis larvae swim with the ciliary band consisting of two tiers of 12 large multiciliated cells (Figs 1a and 2a, b). Staining of all neurites with anti-acetylated-tubulin antibody showed axonal contact between the eyespot and the ring nerve that runs underneath the ciliary band (Fig. 2a). To trace in detail the axonal connections between the eyespots and the ciliated cells, we followed the photoreceptor cell through 70-nm electron microscopy sections. We found that the cell body of the photoreceptor extends a basal axon that joins the ring nerve and forms synapses on the multiciliated cells (Fig. 2b, c and Supplementary Fig. 2). Immunohistochemistry allowed further characterization of the photoreceptor axon. We identified the 7-to-9-amino-acid neuropeptide x(3,5)-FVRIamide (in short, FVRI-amide) as a photoreceptor marker by co-staining the rhabdom microvilli using rhodamine-phalloidine and the photoreceptor cell body using an antibody against FVRI-amide (Fig. 2d). FVRI-amide immunostainings also highlighted the photoreceptor axon along its entire length (Fig. 2e). In agreement with the electron microscopy data, we observed varicose thickenings adjacent to ciliated cells indicative of synaptic contact (Fig. 2e).

To identify the neurotransmitter of the photoreceptor we performed whole-mount *in situ* hybridization with candidate marker genes. Staining for *Platynereis* vesicular acetylcholine transporter²⁰ (*VACHT*) identified a cholinergic neuron in the region of the eyespot (Fig. 2f). Double whole-mount *in situ* hybridization²¹ with *VACHT* and the photoreceptor-specific FVRI-amide neuropeptide precursor probe showed cellular-resolution co-expression²², demonstrating that the photoreceptor is cholinergic (Fig. 2g). We also found that an $\alpha 9/\alpha 10$ -subunit of nicotinic acetylcholine receptor is expressed in the ciliated cells (Fig. 2h). These results indicate that the eyespot photoreceptor works as a sensory-motor neuron that directly regulates phototactic steering by acetylcholine (ACh). To test the involvement of ACh we performed phototaxis experiments in the presence of the ACh receptor antagonists mecamylamine and α -bungarotoxin. The presence of both antagonists obliterated phototaxis without affecting the speed of swimming (Fig. 2i, Supplementary Fig. 3 and Supplementary Movie 3).

¹Developmental Biology Unit, ²Cell Biology & Biophysics Unit, European Molecular Biology Laboratory, Heidelberg 69117, Germany. ³Institut für Biologie, Systematik und Evolution der Tiere, Freie Universität Berlin, Berlin 14195, Germany. †Present address: Max Planck Institute for Developmental Biology, Tübingen 72076, Germany.

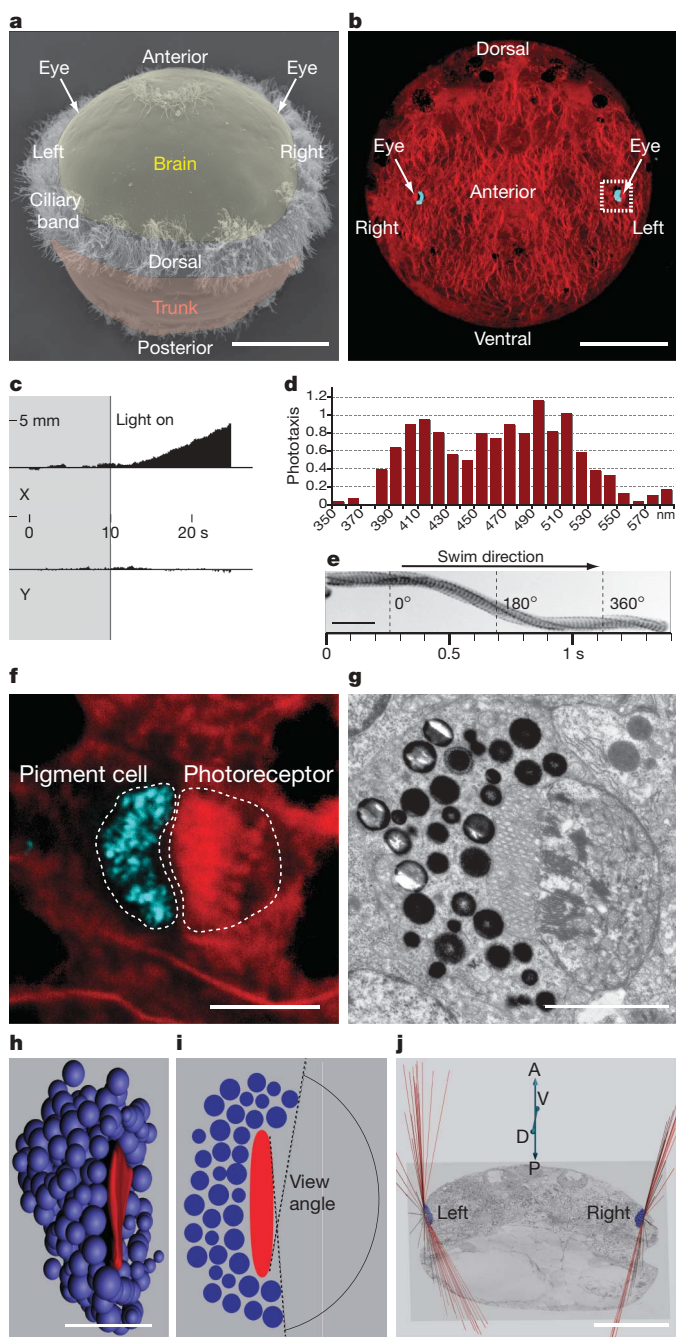


Figure 1 | Larval eyespots mediate phototaxis in *Platynereis*. **a**, Scanning electron micrograph of a 48-h larva. **b**, Anterior view showing pigment autofluorescence (cyan) with the membrane label bodipy564/570-propionic acid (red). Boxed region is enlarged in **f**. **c**, Quantification of average phototactic displacement with directional light after 10 s. Positive phototaxis is indicated as positive X values ($n > 200$ larvae). **d**, Action spectrum of phototaxis. Efficiency is measured as the slope of the displacement plots (compare to panel **c**; $n > 150$ larvae). **e**, Superposition of video frames (60 frames per second (f.p.s.)) of a spiralling phototactic larva, illumination is from the right. The anterior end is forward, the anterior–posterior axis is parallel to the helical axis. The axial rotation is shown. **f**, Pigment (cyan) with bodipy564/570-propionic acid (red), the images is enlarged from **b**. Cell outlines are shown. **g**, Electron microscopy section showing the pigment cup surrounding the photoreceptor rhabdom. **h**, Three-dimensional reconstruction of eyespot on the basis of electron microscopy sections. **i**, Schematic view showing how view angles were determined. **j**, Reconstruction of the view angle of eyespots. The anterior–posterior (A–P) and dorsal–ventral (D–V) axes are indicated. Scale bars, 50 μm (**a**, **b**, **j**), 500 μm (**e**), 5 μm (**f**) and 2 μm (**g**, **h**).

The direct cholinergic innervation of adjacent ciliated cells by the eyespot photoreceptor suggested a direct effect of light perception on ciliary beating. To test this we selectively illuminated one eyespot with a 500 nm monochromatic light beam of 10 μm diameter (Supplementary Fig. 4) and recorded ciliary beating under >650 nm transmission light. On illumination up to five ciliated cells on the side of the eyespot changed their beating pattern with a fast response time of ~ 80 ms (Fig. 3a–c and Supplementary Movie 4). The response was inhibited by mecamylamine, consistent with a direct role for the cholinergic photoreceptor in triggering it (data not shown). The response involved a reduction in ciliary beat frequency (Fig. 3d) and an increase in the amplitude of the recovery stroke (Fig. 3e). No change was observed in either the beat pattern or the beat frequency in cells on the other side of the larva (Fig. 3d and data not shown). The ciliated cells that are closest to the eyespot, and can therefore be directly contacted by the photoreceptor axon, showed the strongest response (Fig. 3c, d). Under repeated stimulation with light/dark cycles (215 ms light, 215 ms dark) the response did not get attenuated (Fig. 3f), showing that the eyespot photoreceptor can quickly dark-adapt and is optimised to detect rapid intensity changes during axial rotation. To understand how the altered stroke pattern affects the forces exerted by cilia we visualized the water flow around immobilised larvae using 1 μm microspheres (Figs 3g and 4a). We measured a decreased flow next to the affected cilia after selective illumination of one eyespot (Fig. 3h; dark speed (\pm s.d.): 0.77 ± 0.09 mm s^{-1} , $n = 36$ beads; light speed: 0.39 ± 0.08 mm s^{-1} , $n = 29$ beads).

These results indicate that there is a functional difference between lateral ‘steering’ cilia that show a strong response, and dorsal–ventral ‘non-steering’ cilia that show no or a weak response. To test this difference experimentally we selectively removed all cilia from selected ciliated cell pairs using laser nanosurgery (Supplementary Movie 7) and performed phototactic turning assays. In this assay phototactic larvae had to turn 180° after a sudden change in the direction of the phototactic light stimulus. When we ablated all cilia from cells 7 and 12, turning efficiency—as characterized by the diameter of the turning curves—was strongly reduced compared to non-ablated larvae or larvae with all cilia removed from cells 4 and 10 (Fig. 3i–k).

Next we wanted to understand the effect of such local change in the beating of the steering cilia on larval swimming and phototaxis. As a prerequisite we further characterized the flow generated by ciliary beating by video recordings. Using microspheres we observed a posteriorly directed flow tangent to the ciliary band (Fig. 4a) representing the main driving force of larval swimming. Recordings from apically oriented larvae showed a dorsal–ventral asymmetry in water flow: the six dorsal cell-pairs propelled microspheres slower than the ventral ones (Fig. 4b). Apically oriented larvae also showed a clockwise water flow around their circumference (data not shown). Taking these parameters into account we then mathematically modelled the swimming of *Platynereis* larvae, considering that they are spheres propelled in a viscous fluid²³ by equatorial cilia that produce dorso-ventrally asymmetric propulsion forces tangent to the sphere and torque that results in axial rotation (Fig. 4c). These parameters are sufficient to generate helical trajectories in a computer simulation of larval swimming (Fig. 4e, <http://www.cytosim.org/platynereis>). Such spiralling is usually the case for self-propelled objects, if the forces applied on the surrounding fluid are asymmetric and constant. In both phototaxis^{24,25} and chemotaxis²⁶, the swimming forces have to be modulated to orient the trajectory towards the desired direction. To model this effect, we introduced local modulation of the ciliary beating forces as a function of the amount of light received by the eyespots. We numerically calculated the resulting trajectories and analysed them (see Supplementary Information for details). The model showed that a local change in ciliary propulsion force is sufficient to effectively steer larval swimming towards the light source (Fig. 4d, e and Supplementary Movies 5 and 6). The trajectories were

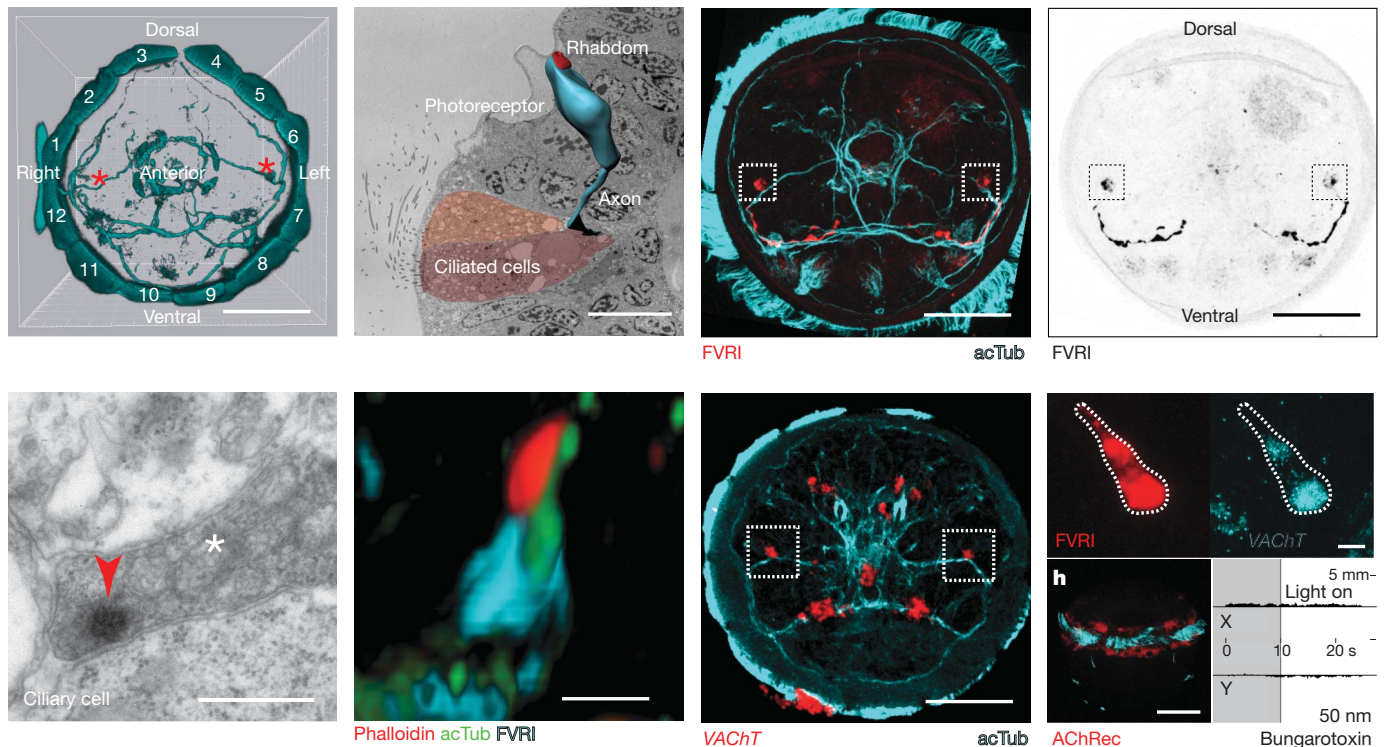


Figure 2 | The eyespot photoreceptor directly regulates phototaxis via ACh. **a**, Anterior view of larva immunostained for acetylated tubulin. Ciliated cells are numbered, the position of the eyespots is marked by red asterisks. **b**, Three-dimensional reconstruction of eyespot photoreceptor (cell body: cyan, rhabdom: red) and axonal contact to ciliated cells (two cells differentially coloured). **c**, Synaptic contact between the photoreceptor axon and a multiciliated cell. The axon is indicated by a white asterisk; the red arrowhead points to a synapse. **d**, Anti-FVRI-amide (cyan), anti-acetylated tubulin (acTub, green) and rhodamine-phalloidin (red) staining of the photoreceptor cell body, dendrite and rhabdom, respectively.

e, Immunostaining for FVRI-amide (red) and acetylated tubulin (cyan). **f**, Whole-mount *in situ* hybridization for VACHT (red), counterstained for acetylated tubulin (cyan). **g**, Double whole-mount *in situ* hybridization for FVRI-amide precursor (red) and VACHT (cyan), the outline of the cell is marked. **h**, Whole-mount *in situ* hybridization for the $\alpha 9/\alpha 10$ -subunit of nicotinic ACh receptor (AChRec) counterstained for acetylated tubulin. **i**, Quantification of phototaxis in the presence of 50 nM α -bungarotoxin ($n > 200$ larvae; compare to Fig. 1c). **d**, **g** and **h** show lateral views, all other panels are anterior views. Scale bars, 50 μ m (**a**, **e**, **f**, **h**), 10 μ m (**b**), 0.5 μ m (**c**) and 5 μ m (**d**, **g**).

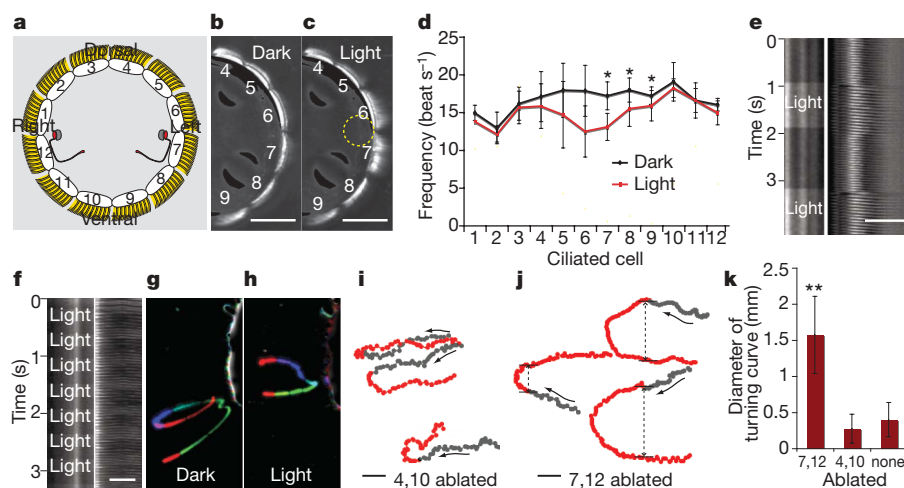


Figure 3 | Light regulation of ciliary beating. **a**, Schematic anterior view of a larva. **b**, **c**, Effect of local illumination on ciliary beating. Superposition of frames from a 60 f.p.s. movie taken in the dark (**b**) or when the left eyespot was illuminated locally (yellow dashed circle) (**c**). A change in beating pattern is apparent in cells 6 and 7 (**c**). **d**, Beat frequencies of cilia in the dark and under selective illumination of the left eyespot. Asterisks, $P < 0.05$ (paired *t*-test), $n = 3$ larvae. **e**, Kymograph from cilia from a laterally oriented larva during selective stimulation of the adjacent eyespot. **f**, Kymograph of cilia upon 215 ms light/dark selective stimulation of the

adjacent eyespot. **g**, **h**, Superposition of frames from a 60 f.p.s. movie showing a microsphere propelled by ciliated cell pair 7 in the dark (**g**) and after illumination of the left eyespot (**h**). **i**, **j**, Phototactic turning trajectories of larvae after 180° reversal of directional light source with laser ablation of cells 4,10 (**i**) and 7,12 (**j**). **k**, Diameter of phototactic turning curves ($n = 8$ larvae each). Double asterisk, $P < 0.0001$ (unpaired *t*-test) between controls and 7,12 ablated larvae. Error bars, s.e.m. (**d**) and s.d. (**k**). Scale bars, 50 μ m (**b**, **c**), 10 μ m (**e**, **f**, **g**) and 100 μ m (**i**, **j**).

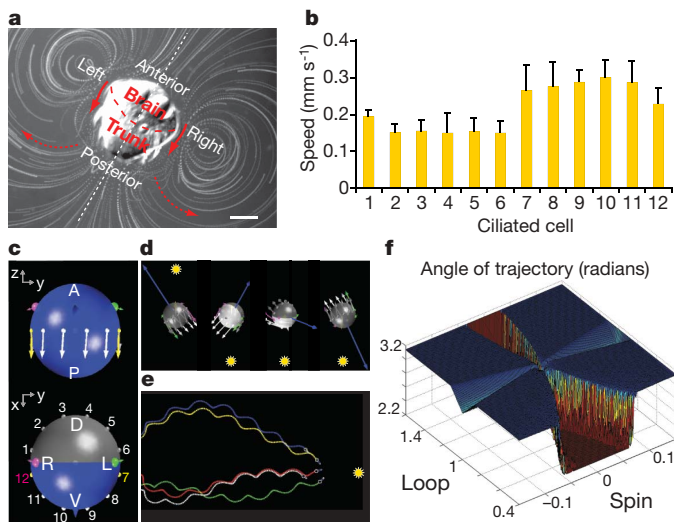


Figure 4 | Computer simulations of phototaxis. **a**, Flow around an immobilised laterally oriented larva. Superposition of video frames showing microspheres carried by cilia-generated flow. Red arrows show flow direction. The ciliary band is indicated with a dashed line. Scale bar, 50 μm . **b**, Quantification of the speed of microspheres propelled by each of the 12 ciliated cell pairs, error bars show s.d., $n > 4$ beads. **c**, Simulated larvae are spherical and propelled by equatorial forces (arrows). The eyes detect light from the direction shown by small arrows and control the cilia 7 and 12. A, anterior; P, posterior; D, dorsal; V, ventral; R, right; L, left. **d**, Simulated reorientation to an infinite light source coming from the opposite direction (yellow dots). **e**, Simulated spiral phototactic trajectories of larvae swimming to a local light source (yellow spot on the right). **f**, Precision of phototaxis as a function of spin and loop. Colour tone indicates the angle (in radians) of the swimming trajectories to the light direction.

similar to the trajectories recorded for *Platynereis* larvae (Figs 1e and 4e).

To gain further insights into the mechanism of phototaxis we used our computer model to test how varying the dorsal–ventral asymmetry in ciliary beating (loop factor), as well as the speed and direction of axial rotation (spin factor), would affect the efficiency of phototaxis. To scan the parameter space we ran 300,000 simulations and plotted the precision of phototactic steering as determined by the angle of the swimming trajectories to the light vector (Fig. 4f). This showed that when axial rotation is not present or is very slow (spin close to 0) phototactic orientation does not occur (Fig. 4f). These larvae only swim in a circle with a plane perpendicular to the light vector. Axial rotation restores the ability to steer in three-dimensions and is therefore essential for phototactic navigation. The plot also showed that straight-swimming larvae (that is, larvae with no dorsal–ventral asymmetry (loop = 1) but performing axial rotation (spin \neq 0)), are imprecise and swim at an angle towards the light rather than directly towards it. This is due to an anterior blind zone in the view field of the larvae within which changes in light direction do not affect the orientation of larvae (compare with Fig. 1j). The spiralling resulting from the dorsal–ventral asymmetry (loop \neq 1) compensates for this and allows the larvae to scan the space in all directions. Our simulations thus reveal that the spiral swimming mode is necessary for precise three-dimensional navigation using simple direction sensing eyespots.

Spiral swimming is the rule among phototactic organisms^{10–14}. It is generally observed in ciliated zooplankton larvae and also occurs in phototactic single-celled algae such as *Chlamydomonas*^{24,25}. Although the swimming mechanics of *Chlamydomonas* is different, the strategy for phototaxis is similar (an eyespot with a restricted view field that rotates in a spiral and directly regulates the beating of cilia). In each case, eyespots and cilia-driven helical swimming are intimately linked functionally and co-adapted for efficient phototaxis. In *Platynereis*, this is reflected by the direct sensory-motor coupling between the

eyespot photoreceptor and the locomotor ciliary cells. In sponge and cnidarian zooplankton larvae, directional light detection and ciliary locomotor output are combined in a single cell^{27,28}. We therefore propose that such direct sensory-motor coupling between a light sensor cell and effector cilia mediating phototaxis represents an ancestral condition in the evolution of animal eyes.

METHODS SUMMARY

Behaviour. *Platynereis* larvae were obtained from an established breeding culture, following that previously described²⁹, and were raised at 18 °C in natural sea water. Behavioural experiments were performed in sea water using 34–52-h-old *Platynereis* larvae. Receptor antagonists were added to the sea water and larvae were incubated for 5 min before recordings. Water flow was highlighted using 1 μm FluoSpheres (Invitrogen).

Laser ablations. Laser ablations of eyespots and cilia were performed on immobilised larvae using either the self-implemented EMBL laser nanosurgery¹⁸ Axiovert 200M microscope (Carl Zeiss), or a CellCut System from Olympus.

Staining. Immunostainings and whole-mount *in situ* hybridization were performed using established protocols^{21,22} on precisely staged larvae. Reflection imaging of nitroblue tetrazolium (NBT)/5-bromo-4-chloro-3-indolyl phosphate (BCIP) stained whole-mount *in situ* hybridization samples was performed as described²².

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 11 September; accepted 27 October 2008.

1. Darwin, C. *On the Origin of Species by Means of Natural Selection: or the Preservation of Favoured Races in the Struggle for Life* (John Murray, 1859).
2. Arendt, D. & Wittbrodt, J. Reconstructing the eyes of Urbilateria. *Phil. Trans. R. Soc. Lond. B* **356**, 1545–1563 (2001).
3. Gehring, W. J. & Ikeo, K. *Pax 6*: mastering eye morphogenesis and eye evolution. *Trends Genet.* **15**, 371–377 (1999).
4. Pichaud, F. & Desplan, C. Pax genes and eye organogenesis. *Curr. Opin. Genet. Dev.* **12**, 430–432 (2002).
5. Thorson, G. Light as an ecological factor in the dispersal and settlement of larvae of marine bottom invertebrates. *Ophelia* **1**, 167–208 (1964).
6. Hays, G. C. A review of the adaptive significance and ecosystem consequences of zooplankton diel vertical migrations. *Hydrobiologia* **503**, 163–170 (2003).
7. Angel, M. V. & Pugh, P. R. Quantification of diel vertical migration by micronektonic taxa in the northeast Atlantic. *Hydrobiologia* **440**, 161–179 (2000).
8. Arendt, D. *et al.* Development of pigment-cup eyes in the polychaete *Platynereis dumerilii* and evolutionary conservation of larval eyes in Bilateria. *Development* **129**, 1143–1154 (2002).
9. Forward, R. B. Diel vertical migration: zooplankton photobiology and behaviour. *Oceanogr. Mar. Biol. Annu. Rev.* **26**, 361–393 (1988).
10. Jennings, H. S. On the significance of the spiral swimming of organisms. *Am. Nat.* **35**, 369–378 (1901).
11. Brandenburger, J. L., Woollacott, R. M. & Eakin, R. M. Fine structure of eyespots in tornarian larvae (Phylum: Hemichordata). *Z. Zellforsch. Mikrosk. Anat.* **142**, 89–102 (1973).
12. Marsden, J. R. Swimming in response to light by larvae of the tropical serpulid *Spirobranchus giganteus*. *Mar. Biol.* **83**, 13–16 (1984).
13. Witman, G. B. *Chlamydomonas* phototaxis. *Trends Cell Biol.* **3**, 403–408 (1993).
14. Maldonado, M. The ecology of the sponge larva. *Can. J. Zool.* **84**, 175–194 (2006).
15. Rhode, B. Development and differentiation of the eye in *Platynereis dumerilii* (Annelida, Polychaeta). *J. Morphol.* **212**, 71–85 (1992).
16. Lacalli, T. C. Structural correlates of photoresponse in trochophore larvae. *Can. J. Zool.* **66**, 1004–1006 (1988).
17. Arendt, D. *et al.* Ciliary photoreceptors with a vertebrate-type opsin in an invertebrate brain. *Science* **306**, 869–871 (2004).
18. Colombelli, J., Grill, S. W. & Stelzer, E. H. K. Ultraviolet diffraction limited nanosurgery of live biological tissues. *Rev. Sci. Instrum.* **75**, 472–478 (2004).
19. Colombelli, J., Reynaud, E. G. & Stelzer, E. H. Investigating relaxation processes in cells and developing organisms: from cell ablation to cytoskeleton nanosurgery. *Methods Cell Biol.* **82**, 267–291 (2007).
20. Denes, A. S. *et al.* Molecular architecture of annelid nerve cord supports common origin of nervous system centralization in Bilateria. *Cell* **129**, 277–288 (2007).
21. Tessmar-Raible, K. *et al.* Fluorescent two-color whole mount *in situ* hybridization in *Platynereis dumerilii* (Polychaeta, Annelida), an emerging marine molecular model for evolution and development. *Biotechniques* **39**, 460–464 (2005).
22. Jékely, G. & Arendt, D. Cellular resolution expression profiling using confocal detection of NBT/BCIP precipitate by reflection microscopy. *Biotechniques* **42**, 751–755 (2007).
23. Denny, M. W. *Air and Water: The Biology and Physics of Life's Media* (Princeton Univ. Press, 1993).
24. Josef, K., Saranak, J. & Foster, K. W. Ciliary behavior of a negatively phototactic *Chlamydomonas reinhardtii*. *Cell Motil. Cytoskeleton* **61**, 97–111 (2005).

25. Josef, K., Saranak, J. & Foster, K. W. Linear systems analysis of the ciliary steering behavior associated with negative-phototaxis in *Chlamydomonas reinhardtii*. *Cell Motil. Cytoskeleton* **63**, 758–777 (2006).
26. Friedrich, B. M. & Jülicher, F. Chemotaxis of sperm cells. *Proc. Natl Acad. Sci. USA* **104**, 13256–13261 (2007).
27. Leys, S. P. & Degnan, B. M. Cytological basis of photoresponsive behavior in a sponge larva. *Biol. Bull.* **201**, 323–338 (2001).
28. Nordström, K., Wallén, R., Seymour, J. & Nilsson, D. A simple visual system without neurons in jellyfish larvae. *Proc. Biol. Sci.* **270**, 2349–2354 (2003).
29. Dorrestijn, A. W. C. *et al.* Molecular specification of cell lines in the embryo of *Platynereis* (Annelida). *Dev. Genes Evol.* **202**, 260–269 (1993).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank K. Miura for advice on behavioural assays and digital image processing, K. Tessmar-Raible for identifying FVRI-amide in our expressed-sequence-tag collection, G. Plickert for advice on antibody production, D. Nilsson for advice on action spectra, and H. Snyman for technical assistance. This work was supported by Center for Modelling and Simulation in the Biosciences

and by grants from Deutsche Forschungsgemeinschaft (HA 4443/1-1 and DA Ar387/1-1). G.J. was supported by a FEBS Long Term Fellowship during the initial phase of the project

Author Contributions G.J. designed the experiments, characterized eyespots, performed behavioural experiments, coordinated the collaborations and wrote the paper. J.C. developed instrumentation for selective eye illumination and performed eyespot and cilia ablations. K.G. characterized FVRI expression. F.N. programmed and performed the simulations and wrote the paper. H.H. provided electron microscopy data and three-dimensional reconstructions. D.A. initiated the project, provided ideas and wrote the paper. All authors discussed the results and commented on the manuscript.

Author Information The sequences for the FVRI-amide neuropeptide precursor and the $\alpha 9/\alpha 10$ subunit of nicotinic ACh receptor were deposited in the GenBank database with the accession numbers FJ358432 and FJ358433, respectively. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to G.J. (gaspar.jekely@tuebingen.mpg.de) or D.A. (arendt@embl.de).

METHODS

Larva fixations. For immunohistochemistry and whole-mount *in situ* hybridization, larvae were fixed in 4% paraformaldehyde in PBS + 0.1% Tween-20 and stored in methanol at -20°C . For transmission electron microscopy, larvae were fixed in 2.5% glutaraldehyde in PBS for 1 h at 4°C with traces of ruthenium red (Sigma), rinsed several times in PBS, and post-fixed with 1% OsO_4 in PBS for 20 min at 4°C . Larvae were dehydrated in a graded acetone series, transferred to propylene oxide and embedded in araldite for 48 h at 60°C . Complete series of silver interfering sections were cut on a Leica EM UC6 microtome, placed on formvar (EMS)-coated copper grids and stained with uranyl acetate and lead citrate. For scanning electron microscopy, larvae were fixed and dehydrated in the same manner, critically point dried in a Balzers CPD 030 (Balzers) and coated with gold in a Balzer Union SCD 040.

Antibody. An FVRI-amide peptide with an amino-terminal Cys was conjugated via the Cys residue to keyhole limpet hemocyanin, and used to immunise rabbits. The antisera were affinity purified on CFVRI-amide coupled to Sulfolink gel (Pierce).

Laser ablations. The Axiovert 200M microscope (Carl Zeiss) and the CellCut System from Olympus were both equipped with a subnanosecond pulsed laser at 355 nm. Eyespot ablations were performed by localized intravital cavitation of the pigment cell induced with 30 nJ energy per pulse (with typically 10 laser pulses). If the larva was oriented upside-down and the laser beam had to focus through the scattering tissues energies of up to 1 μJ were used. Cilia removal was performed by cavitation in sea water at 5 μm distance from the ciliated cell bodies. Energies in the range of 300–500 nJ were used (typically with 20 to 40 pulses along the targeted ciliary band cell), without apparent damage to the cell bodies. Simultaneous imaging of the ablated sample was only possible on the EMBL system with a Hamamatsu ORCA CCD camera.

Microscopy. Confocal images were taken using a Leica TCS SP2 confocal microscope with a $\times 40$ oil-immersion objective using the appropriate laser lines and a pinhole of airy unit 2. Transmission electron microscopy images were taken with a Phillips CM 120 Biotwin equipped with Ditas Digital Imaging Plates. Three-dimensional reconstructions were performed using Autodesk 3ds Max after manual alignment of image stacks. Scanning electron microscopy images were taken with a FEI Quanta 200 microscope. Recordings of phototactic larvae were done using a Leica MZ16 microscope equipped with a Sanyo VCB-3524P camera. Recordings of ciliary beating were done on immobilised larvae using a Zeiss Axiovert microscope equipped with a DMK 21BF04 camera (The Imaging Source). Larvae were immobilised between a slide and a coverslip spaced with adhesive tape.

Image processing. Stacks of confocal images or video recordings were processed using ImageJ 1.38J and Imaris 5.5.1. Contrast was adjusted equally across the entire image. Phototactic movies were analysed using a custom made ImageJ plugin.

Selective eyespot illumination. For selective eyespot illumination we used spectrally selectable monochromatic light (Till Photonics Polychrome II monochromator) coming through an optical guiding fibre of 1 mm diameter and coupled it into the back of the fluorescence port of a Zeiss Axiovert upright microscope. Through the use of a plano-convex lens (focal length 50 mm, Melles Griot) and of an iris, placed 250 mm and 190 mm from the back lens of the port, respectively, we formed a disc of light onto the object plane of a $\times 10$ 0.3 numerical aperture Zeiss objective lens. The diameter of the disc was 10 μm . The aperture of the iris did not exceed 500 μm so that the numerical aperture of the beam at the sample location could be considered infinitely low (straight column of light). To allow axial displacement of the light disc, a manual X-Y mount holding a mirror was placed between the lens and the optical fibre. To perform switchable illumination we developed a software tool to control monochromator wavelength through a RS232 interface.

LETTERS

The zinc-finger protein Zelda is a key activator of the early zygotic genome in *Drosophila*

Hsiao-Lan Liang^{1*}, Chung-Yi Nien^{1*}, Hsiao-Yun Liu¹, Mark M. Metzstein², Nikolai Kirov¹ & Christine Rushlow¹

In all animals, the initial events of embryogenesis are controlled by maternal gene products that are deposited into the developing oocyte. At some point after fertilization, control of embryogenesis is transferred to the zygotic genome in a process called the maternal-to-zygotic transition. During this time, many maternal RNAs are degraded and transcription of zygotic RNAs ensues¹. There is a long-standing question as to which factors regulate these events. The recent findings that microRNAs^{2,3} and Smaug⁴ mediate maternal transcript degradation have shed new light on this aspect of the problem. However, the transcription factor(s) that activate the zygotic genome remain elusive. The discovery that many of the early transcribed genes in *Drosophila* share a *cis*-regulatory heptamer motif, CAGGTAG and related sequences^{5,6}, collectively referred to as TAGteam sites⁵ raised the possibility that a dedicated transcription factor could interact with these sites to activate transcription. Here we report that the zinc-finger protein Zelda (Zld; Zinc-finger early *Drosophila* activator) binds specifically to these sites and is capable of activating transcription in transient transfection assays. Mutant embryos lacking *zld* are defective in cellular blastoderm formation, and fail to activate many genes essential for cellularization, sex determination and pattern formation. Global expression profiling confirmed that Zld has an important role in the activation of the early zygotic genome and suggests that Zld may also regulate maternal RNA degradation during the maternal-to-zygotic transition.

In *Drosophila*, an initial wave of zygotic gene transcription occurs between 1 and 2 h of development during the mitotic cleavage cycles 8–13. This is followed by a major burst of activity between 2 to 3 h of development (cycle 14) when the embryo is undergoing cellular blastoderm formation. Many pre-cellular genes contain TAGteam sites in their upstream regulatory regions including several direct targets of Bicoid, Dorsal and other key regulators of patterning^{5–7}. It has been previously demonstrated⁵ that TAGteam sites are required for the early expression of the dorsoventral gene *zen*, and the sex determination genes *sisB* (also known as *sc*) and *Sxl*. To isolate the TAGteam binding factor, we performed a yeast one-hybrid screen with a 91 base-pair (bp) fragment (Fig. 1a, sequences in upper-case) from the *zen cis*-regulatory region^{8,9} (*zen*(91)), which contains four TAGteam sites⁵ (Fig. 1a in red, the first two are the reverse complement). *zld*, encoded by the X chromosomal gene *CG12701* (also known as *vfl*), was selected as the only candidate of the 11 recovered that had the potential to bind specific DNA sequences because it encoded a protein with six C2H2 zinc fingers (represented as green boxes in Fig. 1b). Oligonucleotides (Fig. 1a, underlined sequences) with different TAGteam sites were tested in gel shift assays with the 357 amino acid carboxy-terminal region of Zld fused to glutathione S-transferase (GST–ZldC; Fig. 1b, stippled region). All oligonucleotides tested formed complexes with GST–ZldC, although with different affinities (Fig. 1c, lanes 1–9),

whereas mutations (Fig. 1a, in purple) in the heptanucleotide sequence abolished binding (Fig. 1c, lanes 10–12). Notably, the site with the strongest affinity, CAGGTAG, is the site most over-represented in regulatory elements of pre-blastoderm genes versus post-blastoderm genes⁵. A plasmid expressing full-length Zld protein promoted transcriptional activation of a *zen*(91)-*lacZ* reporter but not a mutated *zen*(91m)-*lacZ* reporter after co-transfection in *Drosophila* S2 cells (Fig. 1d). Taken together, these data strongly suggest that Zld activates transcription of *zen* and probably other TAGteam-containing genes.

zld transcripts were detected in the germline cells of the ovary (Fig. 2a), in unfertilized eggs (Fig. 2b), and throughout early development (Fig. 2c). Later, *zld* becomes restricted to the nervous system and specific head regions (Fig. 2d), as previously shown¹⁰. To analyse

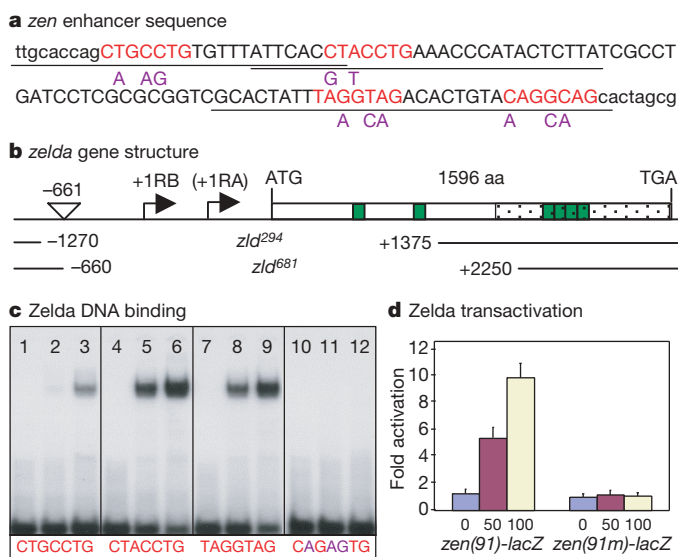


Figure 1 | TAGteam sites bind Zld and mediate transcriptional activation. **a**, DNA sequence of the 91-bp *zen* enhancer (upper-case) plus surrounding sequences (lower-case). Base substitutions are in purple. **b**, Schematic organization of the *zld* locus (*CG12701*; Flybase) with the transcription start sites for the RNA isoforms RB and RA. The P[RS3]UM8171-3 insertion site is between -661 and -660 bp. The nucleotides deleted in *zld*²⁹⁴ and *zld*⁶⁸¹ are indicated as a blank space between solid lines. **c**, Zld binding to oligonucleotides containing different TAGteam sites (denoted beneath each section of the gel). The first lane in each section contains free probe, the second lane contains probe plus 10 ng GST–ZldC, and the third lane contains probe plus 30 ng GST–ZldC. **d**, S2 cells were transfected with 0 ng (blue bar), 50 ng (red bar) or 100 ng (yellow bar) of plasmid expressing *zld* under control of the inducible metallothionein promoter, the *zen*(91)-*lacZ* or *zen*(91m)-*lacZ* reporter plasmids, and the luciferase control. Error bars, s.e.m.; *n* = 3.

¹Department of Biology, New York University, 100 Washington Square East, New York, New York 10003, USA. ²Institute of Human Genetics, University of Utah, Salt Lake City, Utah 84112, USA.

*These authors contributed equally to this work.

zld function, we generated deletion alleles of *zld* by imprecise excision (schematized in Fig. 1b). Hemizygous embryos showed abnormal central nervous system and head development (data not shown), consistent with previous reports of *CG12701* lethal P-insertion phenotypes^{10,11}. *zld* transcripts were not observed in these embryos after cycle 14 (Fig. 2e). However, younger embryos had high levels of maternal *zld* transcripts (data not shown), indicating that maternally loaded *zld* transcripts are degraded during cellularization and replaced with zygotic *zld*.

To eliminate maternal *zld* from embryos, we induced clones of *zld*²⁹⁴ mutant germ cells in the adult female. All resulting embryos were null for maternal *zld* (*M⁻ zld*), and the male embryos were also null for zygotic *zld* (*M⁻ Z⁻ zld*). All early *M⁻ zld* embryos lacked *zld* transcripts (Fig. 2f) but had normal patterns of other maternally deposited factors such as *bicoid* RNAs (Fig. 2g) and the Dorsal protein gradient (data not shown). Unlike *M⁻ Z⁺ zld* embryos, which began to express *zld* ubiquitously in cycle 14 (Fig. 2i), *M⁻ Z⁻ zld* embryos never expressed *zld* (Fig. 2h). However, regardless of their zygotic genotypes, all *M⁻ zld* embryos showed a severely abnormal morphology after cycle 14 (Fig. 2h and j) and did not survive to make cuticle.

Before cycle 14, *M⁻ zld* embryos are similar to wild type (compare Fig. 3a to 3c), except for sporadic nuclear fallout (arrow in Fig. 3c, middle). However, at early cycle 14 the hexagonal-actin network

becomes disorganized and begins to degenerate (Fig. 3d, top) resulting in a multinucleated phenotype (arrow in Fig. 3d, bottom) resembling *nullo*¹² and *Serendipity* α ¹³ (*Sry- α*) mutants. Cellularization does not proceed as furrow canals never move inward (arrow in Fig. 3d, top) like in the wild type (arrow in Fig. 3b, top), and Neurotactin (Nrt) accumulates abnormally in the apical cytoplasm (arrow in Fig. 3h, top)—reminiscent of the *slam* mutant phenotype^{14,15}. Staining with anti-Slam antibody confirmed that Slam protein is mostly absent by mid-cycle 14 (Fig. 3f, top), whereas Slam has moved basally in wild type (Fig. 3e, bottom). In addition, nuclei do not elongate but instead become rounded, enlarged and clump together (see Fig. 3f, and arrow in 3h, bottom). Regions of higher nuclear density were observed (data not shown), a phenotype similar to that obtained by injection of *CG12701* double-stranded RNA¹⁰, which we noticed resembles a *frühstart* (*frs*, also known as *Z600*) phenotype¹⁶. Despite their aberrant morphology, *M⁻ zld* embryos attempt to form a ventral furrow (Supplementary Fig. 1b, c, e) but soon become highly disorganized with only pole cells recognizable (Supplementary Fig. 1f). We rescued the *M⁻ zld* cellularization defects by driving a wild-type copy of *zld* into the germ line using the *ovarian tumour* (*otu*) promoter¹⁷. The cytoskeleton becomes well structured (Fig. 3i, top) and furrow canal ingression is normal (Fig. 3i, bottom) as Slam protein is restored (data not shown).

The broad range of phenotypes strongly indicated that *M⁻ zld* embryos do not express genes essential for cellular blastoderm formation. We assayed the expression of *Sry- α* , *slam* and *nullo*, as well as *sisA*, *sisB*, *sisC* (also known as *os*), *Sxl*, *zen* and *dpp*. None of these genes was activated in *M⁻ zld* embryos (data shown for *Sry- α* , *sisB* and *zen* in Fig. 4b, d and f, respectively), except at the poles in some cases. However, *sna* and *sog*—which are activated by Dorsal¹⁸—were not absent but were delayed in expression by at least two cycles (data not shown), suggesting that Zld facilitates the onset of early gene transcription. Furthermore, the lateral stripes of *sog* were greatly reduced in width (Fig. 4h), indicating that in regions in which Dorsal protein amounts are low, a combinatorial mechanism involving both Dorsal and Zld establishes the broad *sog* domain. Notably, there are two TAGteam sites in the 393 bp *sog* enhancer¹⁹ that lie close to Dorsal binding sites.

Our results indicated that Zld is a global activator of early genes. To test this directly we compared the expression profiles of wild-type and *M⁻ zld* embryos in mitotic cycles 8–13, a time point presumably enriched in genes that are direct Zld targets. One-hundred-and-twenty genes were downregulated and surprisingly 176 genes were upregulated at least twofold ($P \leq 0.05$), in the absence of Zld (Fig. 4k). The downregulated set (Supplementary Table 1) was strongly enriched in genes that are zygotically expressed (Fig. 4l) and involved in early developmental processes (Supplementary Fig. 2), including most of the genes we assayed *in situ* (Fig. 4k). For example, 12 genes involved in cellular blastoderm formation (*nullo*, *slam*, *Sry- α* , *bnk*, *frs*, *btsz*, *halo* and 5 *halo*-like genes²⁰), 6 sex-determination genes (*sisA*, *sisB*, *sisC*, *run*, *Sxl* and *dpp*), and 8 dorsoventral genes (*dpp*, *tld*, *tok*, *tsg*, *tsg-like*, *scw*, *zen* and *zen2*) are in our downregulated data set. Overall, 75% of the early genes previously described as pre-cellular^{5,6,21} are included. This number may be an underestimate because there could be many genes such as *sna* and *sog* that did not make the twofold cutoff (Fig. 4k) but are indeed regulated by Zld.

About 80% of the downregulated genes have TAGteam sites within 2 kilobases (kb) upstream of the transcription start site (Supplementary Table 2), and another 10% have TAGteam sites in introns, such as *slam* with two sites in its first intron, supporting the idea that most of our downregulated genes are direct Zld targets. In addition, the TAGteam sites upstream of the downregulated genes tend to be located very close to the transcription start site within 200 bp (Supplementary Table 2), consistent with the previous finding that early zygotic genes have a statistical over-representation of TAGteam sites close to the start site^{5,6}.

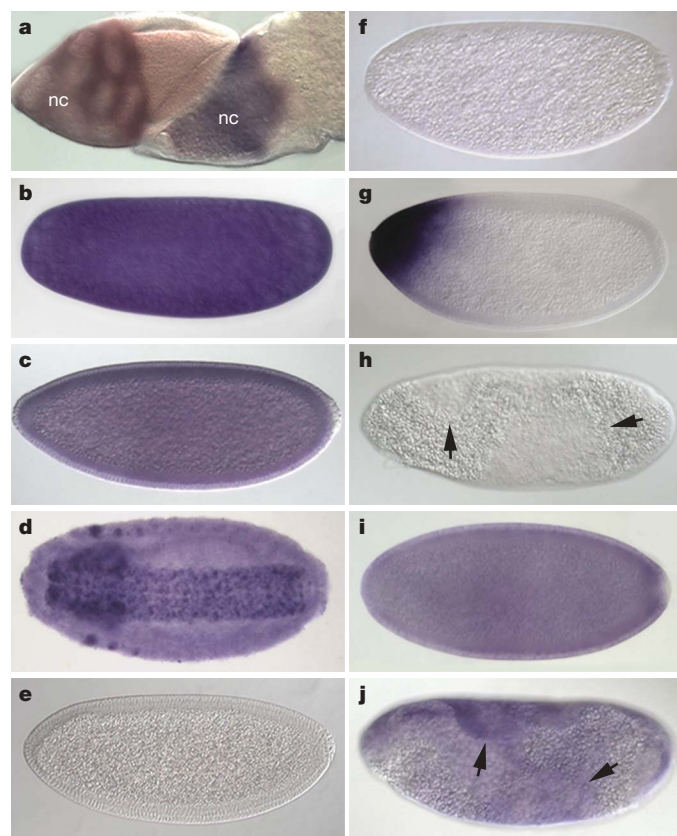


Figure 2 | Maternal *zld* transcripts are lost as zygotic *zld* is activated in cycle 14. **a–j**, Wild-type (**a–d**) and *zld*²⁹⁴ (**e–j**) ovaries (**a**) and embryos (**b–j**) were hybridized with *zld* (**a–f** and **h–j**) or *bcd* (**g**) RNA probes. **a**, Mid-stage (left) and late-stage (right) egg chambers with *zld* transcripts in the nurse cells (nc) but not the columnar follicle cells that overlay the oocyte. **b**, An unfertilized egg is shown. **c**, A cycle 14 embryo undergoing cellularization. **d**, A late-stage embryo is shown. **e**, An *M⁺ Z⁻ zld* cycle 14 embryo showing that maternal *zld* transcripts have disappeared. **f**, An *M⁻ zld* cycle 10–11 embryo is shown. **g**, An *M⁻ zld* cycle 14 embryo has a normal *bcd* pattern. **h**, An *M⁻ Z⁻ zld* late cycle 14 embryo showing anomalous distribution of cytoplasm (arrows). **i**, An *M⁻ Z⁺ zld* early cycle 14 embryo showing onset of zygotic *zld* expression. **j**, An *M⁻ Z⁺ zld* late cycle 14 embryo showing abnormalities (arrows).

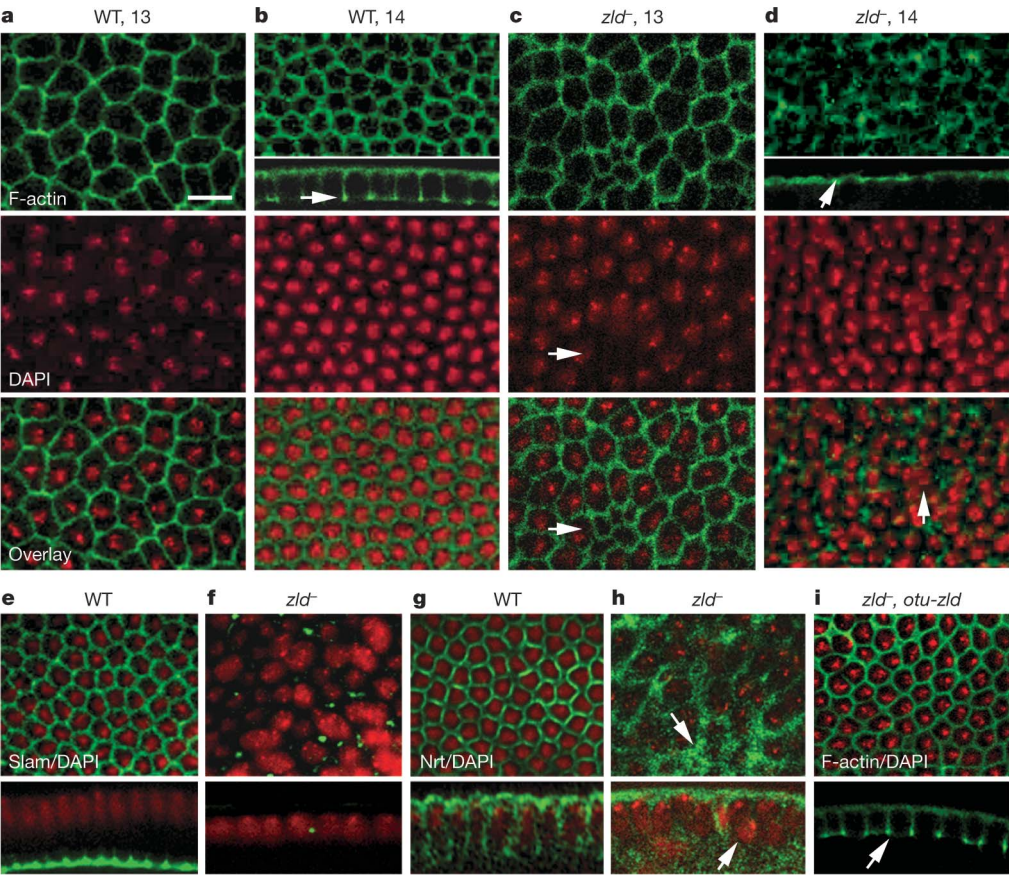
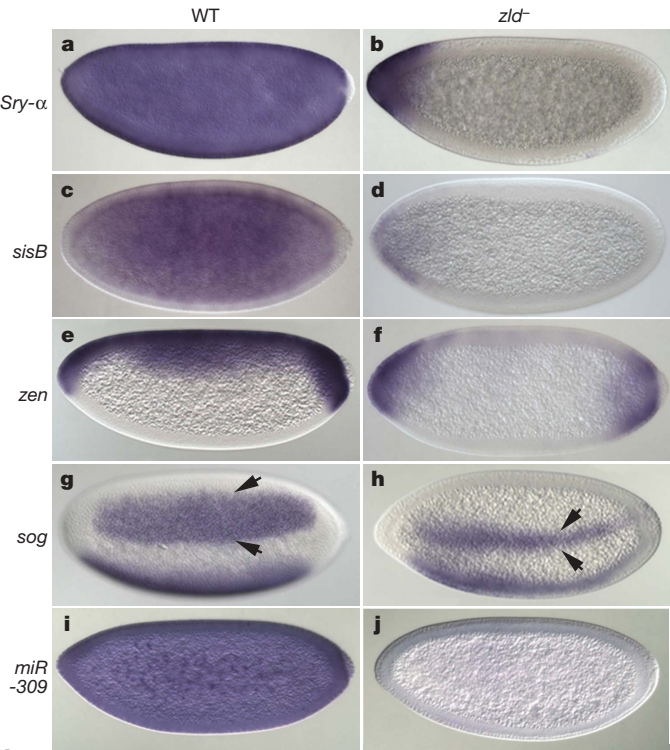


Figure 3 | Maternal *zld* is required for cellularization. **a–i**, Confocal images of wild-type (WT; **a, b, e, g**), *M[−]zld²⁹⁴* (**c, d, f, h**) and rescued (*M[−]zld²⁹⁴; otu-zld*, **i**) embryos (as indicated) stained with phalloidin to detect F-actin, with anti-slam or anti-Nrt antibodies (green), and with DAPI (4,6-diamidino-2-phenylindole) to detect DNA (red). In *M[−]zld* embryos the cytoskeletal network is disorganized and quickly degenerates in early cycle 14 (**d**, top) accompanied by nuclear fallout (**d**, middle). Slam protein disappears in cycle 14 (**f**, top and middle), whereas Nrt accumulates apically (**h**, top, arrow). In *M[−]zld²⁹⁴; otu-zld* embryos the cytoskeleton is organized (**i**, top) and cellularization proceeds (**i**, bottom, arrow). Scale bar, 10 μm.

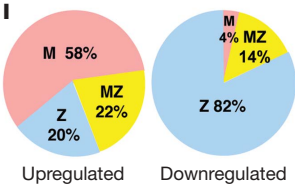


In contrast to the downregulated genes, the upregulated set is strongly enriched in genes that are maternally expressed (Fig. 4l). We considered the possibility that Zld activates components of the RNA degradation machinery that in turn destabilize maternal RNAs. Because the microRNA (miRNA) *miR-309* enhancer²² contains two TAGteam sites, we assayed for *miR-309* primary transcripts in *M[−]zld* embryos (Fig. 4j), and indeed they were absent. It was recently shown that mature *miR-309* miRNAs become abundant during cycle 14, and are involved in maternal transcript turnover in 2–4 h embryos³. Not surprisingly, our 1–2 h (cycles 8–13) data set had no overlap with the 44 published *miR-309* targets³, however 2–4 h profiling experiments should demonstrate whether they are upregulated in the absence of *zld*. We also compared our upregulated genes to those affected by *smaug*⁴, another gene required for the removal of maternally supplied RNAs. We found there was little overlap with the published Smaug targets⁴, suggesting that Zld is involved in a parallel pathway of maternal RNA degradation.

We have demonstrated that Zld functions as a key transcriptional activator during the maternal-to-zygotic transition (MZT) in *Drosophila*. This is the first demonstration of such an activator in any organism. We propose that the biological role of Zld in the pre-blastoderm embryo is to set the stage for vital processes such as cellular blastoderm formation, counting of X chromosomes for dosage compensation and sex determination, and pattern formation, by ensuring

Figure 4 | Zld plays a role in zygotic gene activation and maternal RNA degradation during the MZT. **a–j**, Wild-type (WT; left) and *M[−]zld²⁹⁴* (*zld[−]*; right) mitotic cycle 12–14 embryos were hybridized as indicated. **k**, Summary of expression profiles of 1–2 h wild-type and *M[−]zld²⁹⁴* embryos. Fold ch, fold change with respect to wild type (genes absent in the array data are not included). **l**, Percentage of genes for which there is expression data^{3,5,19} described as maternal (M), zygotic (Z) or both (MZ) in the downregulated (≥ 2 -fold; $n = 105$) and upregulated (≥ 1.5 -fold; $n = 263$) gene sets.

Fold ch	Genes	Genes tested by <i>in situ</i>	Fold ch	Fold ch
≥ 2.0 down	120	<i>sisA</i> 7.2	<i>Sry-α</i> 8.0	
1.5–2 down	159	<i>sisB</i> 4.9	<i>slam</i> 2.0	
Unchanged	6,970	<i>sisC</i> 2.5	<i>nullo</i> 9.2	
1.5–2 up	155	<i>Sxl</i> 3.5	<i>sna</i> 1.8	
≥ 2.0 up	176	<i>zen</i> 3.6	<i>sog</i> 1.2	
Whole array	13,615	<i>dpp</i> 5.7	<i>bcd</i> 1.0	



the coordinated accumulation of batteries of gene products during the MZT. This early preparedness should allow sufficient time for the formation of molecular machines²³ involved in these processes, and so are ready to spring into action during the prolonged interphase of cycle 14.

METHODS SUMMARY

Fly strains. The *zld*²⁹⁴ and *zld*⁶⁸¹ alleles were generated by imprecise excision of the *P{RS3}* (ref. 24) element *UM-8171-3* (FlyBase, Szeged stock centre). The *ovo*^D *FRT19A* stock was generated by transposition of *P{mini w⁺, ovo^{D1}}* (ref. 25) onto *y w sn FRT19A, hsFLP122*. Germline clones were induced in *zld*²⁹⁴ *FRT19A/ovo*^{D1} *FLP122 FRT19A* by the FLP-FRT technique²⁶. Virgin females were collected and mated to *y w, FM7* or *FM7c-ftz-lacZ* males. The *otu-zld* construct was micro-injected into *w¹¹¹⁸* embryos.

Yeast one-hybrid assay. The yeast one-hybrid screen was performed using the Matchmaker One-Hybrid System (Clontech) protocol with the 91-bp *zen* promoter (*zen*(91)) and a 0–6 h *Drosophila* embryonic complementary DNA library fused to the Gal4 activation domain²⁷ (gift from L. Pick).

Molecular biology. DNA binding assays⁹ and *Drosophila* S2 cell transient transfection assays²⁸ were performed as previously described. The fold activation was calculated as a ratio of the normalized (for transfection efficiency) LacZ activity in cells treated with 0.5 mM CuSO₄ and untreated cells.

Analysis of phenotypes. Various RNA probes, antibodies and molecular probes were used to detect gene expression or to visualize the cytoskeleton and nuclei (further described in the Methods). Embryos were viewed by fluorescence microscopy using a Nikon FX-A microscope for whole embryo views, or an Impropvision Yokogawa CSU-10 spinning disk confocal system for grazing and sectional views, and by Nomarski optics using a Zeiss Axiophot microscope.

Microarray analysis. Total RNA was extracted from three independent collections of 1–2 h *y w* and *M⁻ zld* embryos by Trizol (Invitrogen). cDNA was prepared using the GeneChip HT One-Cycle cDNA Synthesis Kit (Invitrogen), labelled with the BioArray HighYield RNA Transcript Labelling Kit (Enzo), and hybridized to Affymetrix *Drosophila* Genome 2 arrays and processed by a GeneChip Fluidics Station 400.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 20 June; accepted 29 August 2008.

Published online 19 October 2008.

- Newport, J. & Kirschner, M. A major developmental transition in early *Xenopus* embryos: II. Control of the onset of transcription. *Cell* **30**, 687–696 (1982).
- Giraldez, A. J. *et al.* Zebrafish miR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* **312**, 75–79 (2006).
- Bushati, N., Stark, A., Brennecke, J. & Cohen, S. Temporal reciprocity of miRNAs and their targets during the maternal-to-zygotic transition in *Drosophila*. *Curr. Biol.* **18**, 501–506 (2008).
- Tadros, W. *et al.* SMAUG is a major regulator of maternal mRNA destabilization in *Drosophila* and its translation is activated by the PAN GU kinase. *Dev. Cell* **12**, 143–155 (2007).
- ten Bosch, J. R., Benavides, J. A. & Cline, T. W. The TAGteam DNA motif controls the timing of *Drosophila* pre-blastoderm transcription. *Development* **133**, 1967–1977 (2006).
- De Renzis, S. D., Elemento, O., Tavazoie, S. & Wieschaus, E. F. Unmasking activation of the zygotic genome using chromosomal deletions in the *Drosophila* embryo. *PLoS Biol.* **5**, 1036–1051 (2007).
- Li, X. *et al.* Transcription factors bind thousands of active and inactive regions in the *Drosophila* blastoderm. *PLoS Biol.* **6**, 365–388 (2008).
- Jiang, J., Rushlow, C. A., Zhou, Q., Small, S. & Levine, M. Individual Dorsal morphogen binding sites mediate activation and repression in the *Drosophila* embryo. *EMBO J.* **11**, 3147–3154 (1992).
- Kirov, N., Zhelnin, L., Shah, J. & Rushlow, C. Conversion of a silencer into an enhancer: evidence for a co-repressor in dorsal-mediated repression in *Drosophila*. *EMBO J.* **12**, 3193–3199 (1993).

- Staudt, N., Fellert, S., Chung, H., Jäckle, H. & Vorbrüggen, G. Mutations of the *Drosophila* zinc finger-encoding gene *vielfältig* impair mitotic cell divisions and cause improper chromosome segregation. *Mol. Biol. Cell* **17**, 2356–2365 (2006).
- Bourbon, H. M. *et al.* A P-insertion screen identifying novel X-linked essential genes in *Drosophila*. *Mech. Dev.* **110**, 71–83 (2002).
- Simpson, L. & Wieschaus, E. F. Zygotic activity of the *nullo* locus is required to stabilize the actin-myosin network during cellularization in *Drosophila*. *Development* **110**, 851–863 (1990).
- Schweisguth, F., Lepesant, J. A. & Vincent, A. The *serendipity alpha* gene encodes a membrane-associated protein required for the cellularization of the *Drosophila* embryo. *Genes Dev.* **4**, 922–931 (1990).
- Lecuit, T., Samanta, R. & Wieschaus, E. *slam* encodes a developmental regulator of polarized membrane growth during cleavage of the *Drosophila* embryo. *Dev. Cell* **2**, 425–436 (2002).
- Stein, J. A., Brohier, H. T., Moor, L. A. & Lehmann, R. Slow as molasses is required for polarized membrane growth and germ cell migration in *Drosophila*. *Development* **129**, 3925–3934 (2002).
- Grosshans, J., Müller, H. & Wieschaus, E. Control of cleavage cycles in *Drosophila* embryos by *frühstart*. *Dev. Cell* **5**, 285–294 (2003).
- Robinson, D. N. & Cooley, L. Examination of the function of two kelch proteins generated by stop codon suppression. *Development* **124**, 1405–1417 (1997).
- Stathopoulos, A. & Levine, M. Genomic regulatory networks and animal development. *Dev. Cell* **9**, 449–462 (2005).
- Markstein, M., Markstein, P., Markstein, V. & Levine, M. S. Genome-wide analysis of clustered Dorsal binding sites identifies putative target genes in the *Drosophila* embryo. *Proc. Natl Acad. Sci. USA* **99**, 763–768 (2002).
- Gross, S. P., Guo, Y., Martinez, J. E. & Welte, M. A. A determinant for directionality of organelle transport in *Drosophila* embryos. *Curr. Biol.* **13**, 1660–1668 (2003).
- Pilot, F., Philippe, J. M., Lemmers, C., Chauvin, J. P. & Lecuit, T. Developmental control of nuclear morphogenesis and anchoring by charleston, identified in a functional genomic screen of *Drosophila* cellularization. *Development* **133**, 711–723 (2006).
- Biemar, F. *et al.* Spatial regulation of microRNA gene expression in the *Drosophila* embryo. *Proc. Natl Acad. Sci. USA* **102**, 15907–15911 (2005).
- Gunsalus, K. C. *et al.* Predictive models of molecular machines involved in *Caenorhabditis elegans* early embryogenesis. *Nature* **436**, 861–865 (2005).
- Ryder, E. *et al.* The DrosDel collection: a set of P-element insertions for generating custom chromosomal aberrations in *Drosophila melanogaster*. *Genetics* **167**, 797–813 (2004).
- Chou, T. B., Noll, E. & Perrimon, N. Autosomal *P[ovo^{D1}]* dominant female-sterile insertions in *Drosophila* and their use in generating germ-line chimeras. *Development* **119**, 1359–1369 (1993).
- Chou, T. B. & Perrimon, N. The autosomal FLP-DFS technique for generating germline mosaics in *Drosophila melanogaster*. *Genetics* **144**, 1673–1679 (1996).
- Yu, Y. *et al.* The nuclear hormone receptor Ftz-F1 is a cofactor for the *Drosophila* homeodomain protein. *Nature* **385**, 552–555 (1997).
- Kirkpatrick, H., Johnson, K. & Laughon, A. Repression of *dpp* targets by binding of *brinker* to *mad* sites. *J. Biol. Chem.* **276**, 18216–18222 (2001).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank the following people for gifts of RNA probes, plasmids, antibodies and fly stocks: J. Erickson, R. Lehmann, R. Cinalli, R. Martinho, C. Navarro, J. Treisman and L. Pick. We thank J. Rhee and A. Chung for help with isolating and characterizing *zld* null mutants, and S. Fu for help locating TAGteam sites in *Zelda* target genes. We are grateful to M. Siegel and K. Birnbaum for help with the microarray analysis, and C. Desplan, P. Struffi, S. Small and R. Lehmann for critical reading of the manuscript. This work was supported by a grant from the National Institutes of Health (GM63024).

Author Information The microarray data have been deposited in NCBI's Gene Expression Omnibus (GEO; <http://www.ncbi.nlm.nih.gov/geo/>) and are accessible through GEO series accession number GSE11231. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to C.R. (chris.rushlow@nyu.edu).

METHODS

Fly strains. The *zld*²⁹⁴ and *zld*⁶⁸¹ alleles were generated by imprecise excision of the *P{RS3}* (ref. 24) element *UM-8171-3* located in 18F2 (Flybase, Szeged stock centre) using a $\Delta 2,3$ source on the CyO chromosome (gift from J. Treisman). These were balanced over *FM7c-ftz-lacZ* (gift from S. Roth). Germline clones were induced in *zld*²⁹⁴ *FRT19A/y w sn P{mini w⁺, ovo^{D1}}* (ref. 25) *FRT19A, hsFLP122* (see later) second or third instar larvae by the FLP-FRT technique²⁶. Virgin females were collected and mated to *y w, FM7/Y* or *FM7c-ftz-lacZ/Y* (to identify the zygotic genotype with respect to *zld*). Escapers (*zld*⁺ germline clone embryos) varied between experiments (1% to 10%). To rescue the *zld* phenotype, one copy of the *otu-zld* transgene on chromosome II (see later) was present in germline clones.

Confirmation of *zld* null alleles. The exact insertion site of *UM-8171-3* was confirmed by sequencing the PCR product of genomic DNA using primers covering the 3' long terminal repeat (LTR) of the P-element and genomic sequences from -1186 to -1162 upstream of the *CG12701* RB start site (Flybase). The P-element inserted between -660 and -661. Deletion break-points of *zld*²⁹⁴ and *zld*⁶⁸¹ were determined by genomic PCR analysis and sequencing of the PCR product using primers outside of the deleted region. *zld*²⁹⁴ deletes sequences -1270 to +1376 and *zld*⁶⁸¹ deletes sequences from -660 to +2250 and leaves part of the P-element. Nucleotide numbering is according to the *Drosophila melanogaster* X chromosome sequence (release v5.1).

Construction of the *ovo*^D *FRT19A* chromosome. We made the *ovo*^D *FRT19A* stock by transposition of *P{mini w⁺, ovo^{D1}}* (ref. 25) onto *y w sn FRT19A, hsFLP122*. In brief, males of genotype *y w sn FRT19A, hsFLP122/Y*; +/ CyO, *A2-3*; *FRT82B ovo^{D1-18}/Sb¹* were crossed to *C(1)DX, y¹ w¹ f¹/Y* females. Individual male offspring that were *w⁺ Cy⁺ Sb* (genotype *y w sn FRT19A, hsFLP122*; */+; *Sb¹ */+*, in which * represents potential *P{mini w⁺, ovo^{D1}}* insertions) were crossed to *C(1)DX, y¹ w¹ f¹/Y* females. Insertions onto the X chromosome were identified as lines in which the *w⁺* was transmitted to all of the sons and to no daughters. These chromosomes were maintained as *y w sn P{mini w⁺, ovo^{D1}} FRT19A, hsFLP122/Y* and *C(1)DX, y¹ w¹ f¹/Y* females. To test for penetrance of the *ovo*^D insertion, and to confirm that the *FRT19A* or *hsFLP122* elements were not inadvertently mobilized, we crossed the *y w sn P{mini w⁺, ovo^{D1}} FRT19A, hsFLP122* males to *y w* and *y w FRT19A* females at 25 °C. We subjected a brood from each of these crosses (24–48-h-old larvae) to a heat shock at 37 °C for 1 h, whereas other broods were not heat shocked. Female offspring were then assayed for fertility and egg laying. We selected a line (4.1) in which *y w sn P{mini w⁺, ovo^{D1}} FRT19A, hsFLP122/y w*, either with or without heat shock, showed complete sterility and laid almost no eggs, whereas the *y w sn P{mini w⁺, ovo^{D1}} FRT19A, hsFLP122/y w FRT19A* showed almost complete sterility in the absence of heat shock, but showed good fertility after heat shock. The male offspring from these fertile females were all *w sn⁺*, indicating that they contained the *y w FRT19A* and not the *y w sn P{mini w⁺, ovo^{D1}} FRT19A, hsFLP122* chromosome, as would be expected if the *P{ovoD}* was dominantly blocking female germline development.

***zld* rescue construct.** The *zld* rescue construct was made by subcloning the full-length *zld* coding region +642 to +6345 relative to the RB +1 start site into the EcoRI site of the pCOG plasmid, which lies between the *otu* promoter and the *K103'* UTR¹⁷ (gift from C. Navarro). Full-length *zld* was prepared by PCR from genomic DNA (Clontech), cloned into the pCR2.1-TOPO vector (Invitrogen), and verified by sequencing.

Site-directed mutagenesis. The following *zen* promoter fragments were subcloned into the EcoRI site of the pCaSpeRhp43-lacZ transformation vector (gift from M. Frasch): two tandem copies of the *zen(91)* fragment (Fig. 1a in upper-case) prepared by PCR of subcloned *zen* genomic DNA, two copies of a mutated version of the *zen(91)* fragment (*zen(91m)*) in which the base substitutions shown in Fig. 1a (in purple) were introduced by PCR site-directed mutagenesis.

In situ hybridization. Embryos were hybridized with digoxigenin-UTP (Roche Biochemicals) RNA probes synthesized from cloned cDNA sequences, or in the case of *miR-309*, primary transcript sequences spanning the cluster²². *lacZ* staining

indicated embryos that harboured the *FM7-ftz-lacZ* chromosome and were thus *M⁺ Z⁺ zld*. Stained embryos were mounted in aquamount (Polysciences) or embedded in araldite (Polysciences). After hybridization, embryos were stained with DAPI (Sigma) to determine the mitotic cycle. Embryos were visualized by fluorescence microscopy using a Nikon FX-A microscope and by Nomarski optics using a Zeiss Axiophot microscope.

Antibody staining. Dilution of antibodies was as follows: mouse anti-Nrt (Developmental Studies Hybridoma Bank) 1:50; rabbit anti-Slam (gift from R. Lehmann) 1:200; rat anti-Dorsal 1:50; rabbit anti-Vasa 1:500. F-actin and DNA were visualized by TRITC-labelled phalloidin (Sigma) and by DAPI (Sigma), respectively. Embryos were viewed by fluorescence microscopy using a Nikon FX-A microscope for whole embryo views, or an Improvision Yokogawa CSU-10 spinning disk confocal system for grazing and sectional views. Images were prepared using Velocity, ImageJ (W. S. Rasband, <http://rsb.info.nih.gov/ij/>), and Adobe Photoshop software.

Yeast one-hybrid assay. The yeast one-hybrid screen was carried out using the Matchmaker One-Hybrid System (Clontech). The *zen* 91-bp fragment (*zen(91)*) with four TAGteam sites was amplified by PCR and cloned into the EcoRI site of the target-reporter vectors (pLacI and pHisi-1). The TAGteam reporter vectors were integrated into the yeast YM4271 strain to generate the yeast reporter strain YM4271[TAG-lacZ, TAG-his]. The YM4271[TAG-lacZ, TAG-his] reporter strain was used to screen a 0–6 h embryonic cDNA library fused to the Gal4 activation domain²⁷. All 121 His3-positive clones from the first screen were subjected to β -galactosidase activity assays for the second screening, and the final 34 positive clones were sequenced to identify the candidates.

Transient transfection assays. The wild-type and mutated *zen* promoter fragments described above were subcloned into the EcoRI site of the pCaSpeRhp43-lacZ reporter vector²⁸ (gift from A. Laughon). The full-length *zld* coding region was cloned between the KpnI and the XhoI sites of pMT/V5-His B expression vector (Invitrogen). *Drosophila* S2 cells were grown at 28 °C in Schneider's medium (Invitrogen) supplemented with 10% FCS. Three million cells were transfected using Effectene Transfection reagent (Qiagen) with a 250 ng plasmid DNA mix containing 100 ng reporter plasmid, 50 ng plasmid constitutively expressing firefly luciferase, expression plasmid as indicated in Fig. 1d and the plasmid pcDNA3 to bring the total amount of DNA to 250 ng. The expression of Zld protein was induced 24 h after transfection by adding 0.5 mM CuSO₄ directly to the medium. The cells were lysed after 24 h and the β -galactosidase and luciferase activities were assayed (Promega) according to the manufacturer's protocols. The fold activation was calculated as a ratio of the normalized (for transfection efficiency) LacZ activity in cells treated with 0.5 mM CuSO₄ and untreated cells.

DNA binding assays. Electrophoretic mobility shift assays were performed as previously described⁹ using the affinity purified C-terminal part of the Zld protein containing four zinc fingers (amino acids 1240–1470) fused to GST (GST-ZldC) and 26 bp oligonucleotide probes overlapping TAGteam sites from the *zen* promoter, except that EDTA was omitted from the binding buffer and ZnSO₄ was added to 10 μ M. The incubation reactions contained 0.1 ng of ³²P-labelled oligonucleotides (see Fig. 1a for nucleotide sequences) and varying amounts of recombinant GST-ZldC protein.

Microarray analysis. Total RNA was extracted from three independent collections of 1–2 h *y w* and *M⁺ zld* embryos by Trizol (Invitrogen). A portion of the collected embryos was fixed and stained with DAPI; 90% were in nuclear cycles 8 to 13. cDNA was prepared using the GeneChip HT one-cycle cDNA synthesis kit (Invitrogen for Affymetrix) and labelled with the BioArray HighYield RNA transcript labelling kit (Enzo). Labelled probes were hybridized to Affymetrix *Drosophila* Genome 2 arrays and processed by a GeneChip Fluidics Station 400. Data were acquired by a GeneChip Scanner 3000 and processed/normalized by Affymetrix GeneChip Operating Software. Genes were identified as present when at least two of the three replicates had present (P) assignment ($P \leq 0.05$). A Student's two-tailed *t*-test analysis was performed on the data from the three biological replicates.

LETTERS

Concerted multi-pronged attack by calpastatin to occlude the catalytic cleft of heterodimeric calpains

Tudor Moldoveanu¹, Kalle Gehring² & Douglas R. Green¹

The Ca^{2+} -dependent cysteine proteases, calpains, regulate cell migration¹, cell death², insulin secretion³, synaptic function⁴ and muscle homeostasis⁵. Their endogenous inhibitor, calpastatin, consists of four inhibitory repeats, each of which neutralizes an activated calpain with exquisite specificity and potency⁶. Despite the physiological importance of this interaction, the structural basis of calpain inhibition by calpastatin is unknown⁷. Here we report the 3.0 Å structure of Ca^{2+} -bound m-calpain in complex with the first calpastatin repeat, both from rat, revealing the mechanism of exclusive specificity. The structure highlights the complexity of calpain activation by Ca^{2+} , illustrating key residues in a peripheral domain that serve to stabilize the protease core on Ca^{2+} binding. Fully activated calpain binds ten Ca^{2+} atoms, resulting in several conformational changes allowing recognition by calpastatin. Calpain inhibition is mediated by the intimate contact with three critical regions of calpastatin. Two regions target the penta-EF-hand domains of calpain and the third occupies the substrate-binding cleft, projecting a loop around the active site thiol to evade proteolysis.

We determined the crystal structure of the complex between m-calpain and residues 134–219 of calpastatin inhibitory repeat 1 (referred to as calpastatin; Fig. 1a, Supplementary Fig. 1 and Supplementary Table 1). The m-calpain heterodimer consists of an 80 kilodalton (kDa) catalytic subunit and a 28 kDa regulatory subunit⁸. The large subunit contains the Ca^{2+} -dependent protease core domain I–II (DI–II)⁷, DIII (which resembles C2 domains involved in membrane targeting) and the Ca^{2+} -binding penta-EF-hand DIV to heterodimerize the homologous DVI of the regulatory subunit (Supplementary Fig. 1b)⁹. The catalytically inactive C105S 80 kDa subunit was used to overcome unwanted proteolysis observed with the wild-type protein¹⁰. The regulatory subunit was substituted with the 21 kDa DVI (Supplementary Fig. 1b)¹⁰. The 72 kDa calpastatin inhibits heterodimeric calpains with nanomolar affinity, being composed of a non-inhibitory 12 kDa leader L domain and four 15-kDa calpain inhibitory repeats (Supplementary Fig. 1b)⁶. Each repeat contains three regions (A–C) predicted to interact with calpain⁶. We optimized the calpain–calpastatin complex by truncating calpastatin from a longer construct (residues 119–238) on the basis of results from limited proteolysis and NMR spectroscopy, both of which identified the amino- and carboxy-terminal mobile regions that impeded crystallization (Fig. 1b and Supplementary Fig. 2).

Calpastatin recognizes the Ca^{2+} -induced conformation of m-calpain but does not coordinate Ca^{2+} in the complex (Fig. 1a). Regions A and C fold as amphipathic helices when bound to the Ca^{2+} -induced hydrophobic pockets in the corresponding penta-EF-hand domains (Supplementary Fig. 3). Previously, homology modelling based on the Ca^{2+} -dependent complex between DVI and region C of calpastatin predicted the mode of binding for region A¹¹. Regions A

and B of calpastatin engage sites in DIV and DI–III, respectively. Region B associates with DI–III to obstruct the active site in the extended substrate-like orientation. Intrinsic disorder in the free state enables calpastatin to adapt structurally on binding to the substrate-binding cleft of calpain, distantly resembling the inhibitory conformation of the broad-specificity, structured protease inhibitors, the cystatins¹² (Supplementary Fig. 4). Region B is anchored on either side of the active site C105S and avoids proteolysis by forcing a kink (Gly 174-Ile-Lys-Glu-Gly 178) between the flanking residues—Leu 173 at the substrate binding S2 subsite and Thr 179 at S1' (Fig. 2a, b). The unprimed and primed substrate binding subsites are N- and C-terminal, respectively, from the scissile bond P1–P1'. S and P distinguish protein and substrate subsites, respectively. N-terminal to the kink, residues Val 161–Leu 170 participate in hydrogen bonds mediated by backbone atoms (Supplementary Table 2) and hydrophobic interactions with the DIII surface that juxtaposes the active site (Fig. 2b). The remainder of region B binds to the DI–II (Fig. 2a). Ala 172 marks the S3 site whereas the Leu 173 interaction is of particular importance at the S2 pocket, the main specificity determinant of calpains^{13,14}. The side chain of Leu 173 at S2 is superimposable to Leu 2 of leupeptin in complex with the protease core DI–II of μ -calpain, μ I–II (ref. 15; Supplementary Fig. 5). C-terminal to the kink, region B engages the S1'–S2'–S3' subsites with residues Thr 179–Ile-Pro 181. The conformation of Pro 181–Pro-Glu-Tyr 184 changes the direction as the Glu 183–Lys 190 helix targets DI.

We tested the mechanism of inhibition by shortening the kink through deletion of Lys 176, Glu 177 or both. In all instances, the low nanomolar half-maximal inhibitory concentration (IC_{50}) values for m-calpain inhibition, derived from initial rate analysis, did not change significantly compared to wild type (Supplementary Fig. 6a). However, all mutants succumbed to proteolysis within the kink, permitting catalytic cleft access and resulting in complete auto-proteolysis of the complex within hours (Fig. 2c). The 5-residue kink is therefore essential to overcome proteolysis and its length has indeed been conserved (Supplementary Fig. 2b)⁶. An 18-residue peptide B1, specific for DI–II (Ala 172–Glu 189), inhibited m-calpain with an IC_{50} of 410 ± 80 nM (mean \pm s.e.m., Supplementary Fig. 6b). Others have shown that a 27-residue peptide, corresponding to Asp 163–Glu 189 of region B, inhibited μ -calpain with an IC_{50} of ~ 30 nM¹⁶. Residues Leu 173–Gly 174 and Thr 179–Ile-Pro 181 corresponding to the substrate subsites P2-distorted P1 and P1'–P2'–P3', respectively, were identified as 'hotspots' that impaired inhibition when replaced with Ala (ref. 17). Replacing the intact 27-residue peptide with the corresponding N- and C-terminal peptides—Asp 163–Gly 174 and Lys 175–Ala 189 (human sequence)—abolished calpain inhibition¹⁷. Taken together, these data suggest that DIII anchoring by region B enhances inhibitory activity >10 -fold and support the structure-based occluding-loop mechanism. The

¹Department of Immunology, St Jude Children's Research Hospital, 332 N Lauderdale, Memphis, Tennessee 38105, USA. ²Department of Biochemistry, McGill University, 3655 Promenade Sir William Osler, Montreal, Quebec H3G 1Y6, Canada.

Ca^{2+} -dependent reversible interaction between calpastatin and calpain is biologically relevant and, in light of the length of the occluding loop, it indicates that the wild-type inhibitor, unlike the shorter loop mutants, may recycle once dissociated from the protease. This overcomes the need for new synthesis of calpastatin to maintain inhibition under conditions of fluctuating Ca^{2+} levels.

Only regions A, C and the DI–II-binding residues of B are conserved among calpastatins (Supplementary Fig. 2b). The divergent intervening sequences connecting regions A, B and C are devoid of electron density (red dots in Fig. 1a). NMR analysis of the complex between ^{15}N -labelled calpastatin and unlabelled calpain (Fig. 1b and Supplementary Fig. 2a) corroborated that the intervening sequences of calpastatin are intrinsically disordered (Supplementary Fig. 2b). Conversely, calpastatin regions A, B and C bind calpain, become ordered and tumble slowly in solution as part of the 111 kDa

complex, and were undetectable by NMR. DIII contains hotspots for proteolysis in the free and Ca^{2+} -bound m-calpain¹⁸, and its extensive interaction with calpastatin in the complex supports its resistance to trypsin (Supplementary Fig. 2c). The modular organization of calpastatin induced on binding to calpain emphasizes the role of the interspersed, flexible/disordered segments in the folding-binding transitions of the structured regions at distant sites in calpain. Similar disorder was confirmed by NMR for a complex between m-calpain and repeat 1 of calpastatin at 10 μM CaCl_2 (ref. 19). Calpastatin regions A and C interact with calpain whereas the intervening regions are disordered¹⁹. Notably, calpastatin region B corresponding to Lys 176–Leu 188, which in our complex targets mainly DI, also interacts with calpain. At this CaCl_2 level the protease core is not aligned for catalysis and, therefore, the N terminus of

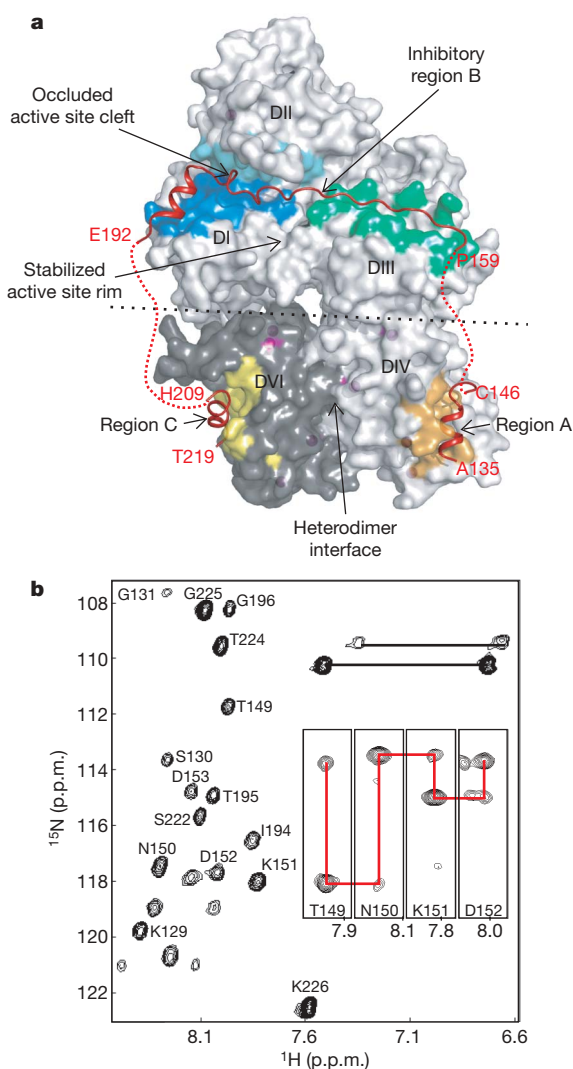


Figure 1 | Complex between Ca^{2+} -bound m-calpain and calpastatin. **a**, Overall structure shows regions A, B and C of calpastatin bound to DIV, DI–III and DVI of calpain, respectively. The intervening sequences of calpastatin are devoid of electron density (red dots). The central part of the inhibitory region B forms the occluding loop at the active site. The active site in the protease core DI–II is stabilized by DIII. Calpain heterodimerization is largely defined at the DIV–DVI interface. Alternate conformations at the interface between the DI–III core and the DIV–DVI heterodimer (black dots) are possible³⁰. **b**, ^{15}N , ^1H -HSQC (heteronuclear single quantum correlation) spectrum of the complex between ^{13}C , ^{15}N -labelled calpastatin (residues 128–226, Supplementary Fig. 2b) and unlabelled calpain identified flexible/disordered residues of calpastatin. Connected sample strips from the HNCA NMR experiment are inset.

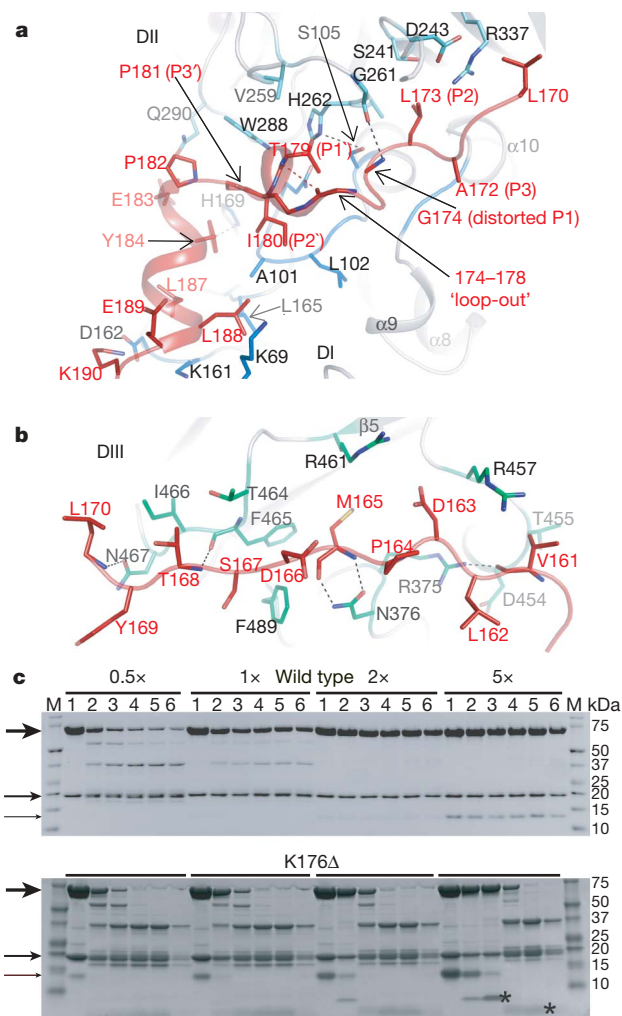


Figure 2 | Binding and inhibition of calpain by region B of calpastatin. **a**, Detailed view of the interaction between calpastatin and the catalytic cleft in the protease core DI–II. At the subsite P1 calpastatin distorts from the substrate path and projects residues 174–178, which kink between the P2 and P1' anchor sites. **b**, Detailed view of the interaction between calpastatin and a surface-accessible groove in DIII. Hydrogen bonds are represented as dashed lines. **c**, Autolysis of m-calpain (10 μM) in the presence of 1 mM CaCl_2 and wild-type calpastatin or the K176 Δ mutant was monitored at 0, 5, 20, 80, 320 min, and 24 h (lanes 1–6, respectively) by SDS–polyacrylamide gel electrophoresis. Calpastatin:calpain (I:E) molar ratios and the size of the molecular markers (lane M) are indicated above and to the left of the panel, respectively. Unprocessed calpain subunits and calpastatin are identified (arrows). Autolysis is blocked by wild-type calpastatin above I:E of 2. The K176 Δ mutant fails to inhibit autolysis even at higher I:E, being digested in the loop to produce Coomassie-stained fragments (asterisks). Fragmentation of the catalytic subunit is complete after ~ 1 h.

region B does not contact the unprimed side of the active site nor the DIII. The disorder in the calpain-bound calpastatin prompted us to predict minimal global conformational changes in calpain instigated by calpastatin binding. This suggests that the rearrangement of calpain domains is mainly induced by Ca^{2+} . Local calpastatin-induced conformational changes in calpain are predicted for the gating loops of the protease core, found in alternative conformations in the structure of $\mu\text{I-II}$ free or bound to inhibitors (Supplementary Figs 5 and 7)^{7,15}.

The calpain–calpastatin structure offers an unprecedented opportunity to study the Ca^{2+} -bound conformation of m-calpain (Fig. 3a). The structures of inactive calpain^{9,20–22}, the Ca^{2+} -bound protease core^{7,23,24} and Ca^{2+} -bound and free DVI (refs 25, 26) have generated valuable, incomplete models for calpain activation⁸. The calpain–calpastatin structure (Fig. 1a) represents the Ca^{2+} -activated conformation of m-calpain revealed by the realignment for catalysis of the protease core DI–II by two Ca^{2+} atoms, as described for $\mu\text{I-II}$ (ref. 7; Supplementary Fig. 7 and Supplementary Table 3). DI–II is intimately associated with DIII, which undergoes conformational changes to interact specifically with DI, serving to stabilize the protease core and to maximize its catalytic activity. The EF-hand DIV and DVI bind four Ca^{2+} atoms each²⁷, mediate the heterodimer interface by pairing of EF-hand 5 as in the apo-calpain⁹ and show small changes between the Ca^{2+} -bound and free conformation (Supplementary Fig. 8 and Supplementary Table 3). Of significance is the Ca^{2+} -induced displacement from DVI of the N-terminal anchor peptide⁹, which is unstructured in the complex.

To investigate the Ca^{2+} -induced conformational changes leading to the heterodimer activation (Supplementary Video), we aligned the Ca^{2+} -bound and free structure²⁰ by overlapping DIII, which provides a central scaffold for the (re)arrangements of the vicinal domains through protein–protein interactions (Fig. 3a). On binding Ca^{2+} , the upper DI–II lobe moves dorsal to frontal with respect to DIII, whereas the lower DIV–DVI lobe moves in the opposite direction. The tension on either side of the protease core, postulated in light of

the Ca^{2+} -free m-calpain structure⁹ and confirmed extensively biochemically⁸, is overcome by Ca^{2+} binding (Fig. 3a). The discovery of the active conformation of the DI–III ensemble is significant as it identifies the missing conserved features at the extensive DI–II–DIII interface (Fig. 3b).

We extrapolated the importance of this interface from structural analysis of the isolated protease core DI–II of m-calpain, mI–II. Owing to intrinsic instability of residues Gly 197–Gly 210 in DI, the unprimed side of the active site in mI–II collapses diminishing activity >1,000-fold compared to full-length m-calpain²⁴. In the Ca^{2+} -bound heterodimer, residues 197–210 are stabilized, through salt bridges and hydrogen bonds, by conserved basic residues in DIII (Fig. 3b and Supplementary Fig. 10). In particular, Arg 417 and Arg 420 from the basic loop, which adopts a different conformation in Ca^{2+} -free calpain (Supplementary Fig. 9), and Arg 469 and Arg 500 from the β -sandwich core of DIII may provide critical support for the labile active site (Fig. 3b).

In limb-girdle muscular dystrophy (LGMD)-2A patients, p94 (calpain 3) point mutants in DIII result in the typical atrophic phenotypes associated with impaired p94 activity in limb-girdle and trunk muscles⁵. The positions 490, 493, 541 and 572 in p94, corresponding to the conserved Arg residues (417, 420, 469 and 500, respectively) in DIII of m-calpain, are mutated to Trp, Gln or Pro in both familial and sporadic forms of the disease (<http://www.dmd.nl>)²⁸. We probed the DI–II–DIII active interface by mutagenesis in m-calpain (Fig. 3c and Supplementary Fig. 11). The R417A and R420A substitutions decreased m-calpain activity to one-half and one-quarter, respectively, and doubled the Ca^{2+} requirement in the latter. The R469A decreased activity >60-fold and doubled the Ca^{2+} requirement. Conservative Lys substitutions at 417, 420 or 469 did not rescue the phenotypes of Ala mutations (Supplementary Fig. 11), collectively underscoring the importance of this interface for sustaining maximal calpain activity and providing an explanation for the effect of disease-causing p94 substitutions in LGMD-2A patients.

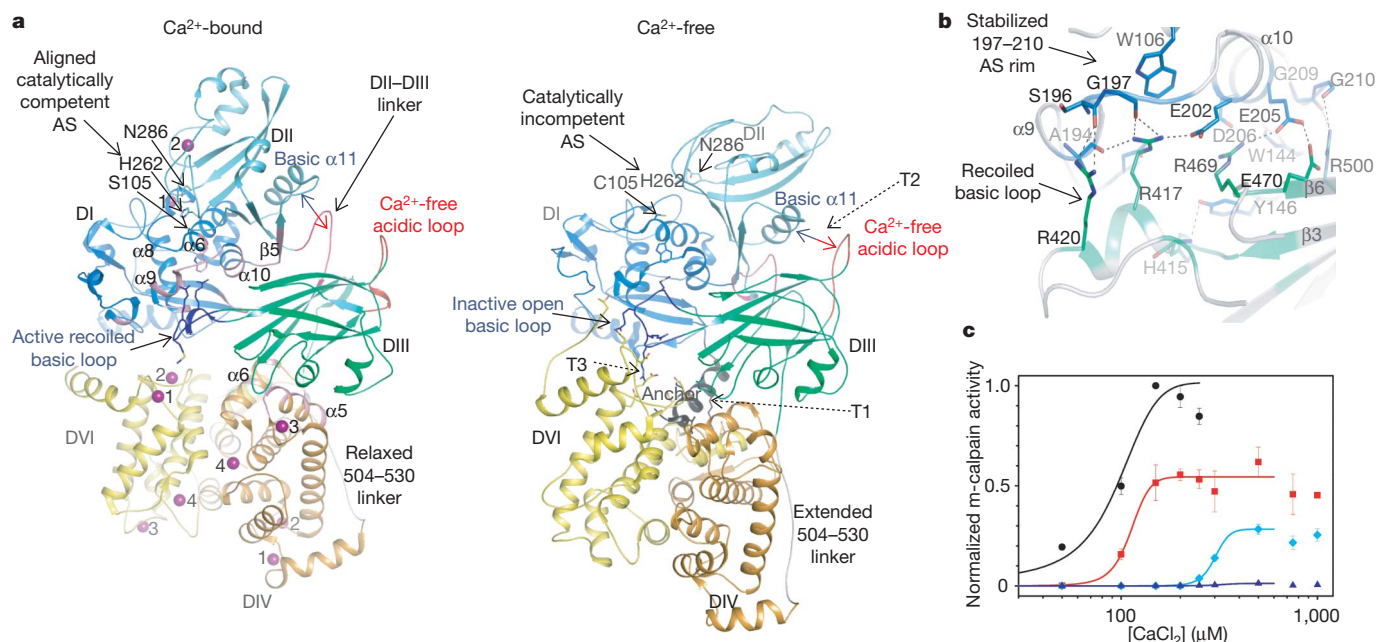


Figure 3 | Stabilization of active Ca^{2+} -bound calpain by DIII. **a**, The Ca^{2+} -bound and free calpain (PDB: 1KFU)²⁰ were aligned structurally based on DIII. The Ca^{2+} -bound DI–II is freed from inactivating tension at the interfaces between the anchor and DVI (T1), the acidic loop of DIII and the basic helix of DII (T2), and the basic loop of DIII and DVI (T3)^{9,21}. In Ca^{2+} -bound calpain, the DIII–DVI linker is relaxed by shedding region 505–513 from the DIII core. DIII is Ca^{2+} -free and contacts DI–II and DIV (pink areas) extensively, but not DVI. The basic region in DIII stabilizes the labile

active site (AS), and the DII–DIII linker contacts the basic helix in DII. Numbering of helices and strands is according to Supplementary Fig. 10. **b**, Detailed view of the interaction between the basic region in DIII and the active site. **c**, Ala substitutions in DIII at Arg 417 (red), Arg 420 (light blue) and Arg 469 (dark blue) impaired m-calpain activity, normalized with respect to the wild type (black); for the latter two it doubled the Ca^{2+} requirement for half-maximal activity. Values are expressed as mean \pm s.d.

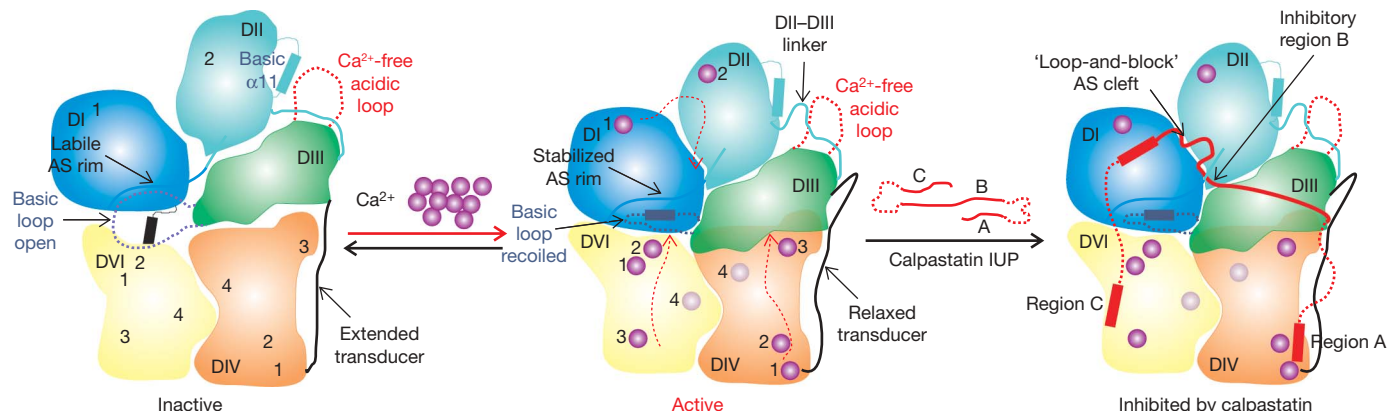


Figure 4 | Calpain-calpastatin proteolytic system. A schematic diagram illustrating the Ca^{2+} -induced activation of calpain and its inhibition by calpastatin. DIII has a fundamental role in relaying the Ca^{2+} -induced structural changes (red dotted arrows) from the peripheral domains to the

catalytically competent yet labile protease core. Concerted binding of the intrinsically unstructured protein (IUP) calpastatin to peripheral domains and the active site of calpain results in low-nanomolar inhibition.

Calpastatin inhibits m- and μ -calpains⁶, which share the 28 kDa regulatory subunit predicted in both to bind calpastatin region C similarly (Supplementary Figs 3b and 12). Modelling of the complex between μ -calpain and calpastatin illustrated that calpastatin regions A and B bind conserved pockets of the 80 kDa catalytic subunit and probably produce a similar set of inhibitory interactions with μ -calpain (Supplementary Fig. 12). We reported that μ I–II is not inhibited by calpastatin repeat 1 (ref. 7). Here we show that trimming down calpastatin repeat 1 from either end to the 18-residue peptide B1 (Ala 172–Glu 189) increased inhibitory potency for μ I–II. The 86-residue repeat 1 (134–219) reduced μ I–II activity with an IC_{50} of $154.7 \pm 16.3 \mu\text{M}$, whereas the 57-residue regions BC (Asp 163–Thr 219) and AB (Ala 144–Lys 190, Fig. 1d), and the peptide B1 inhibited μ I–II with IC_{50} values of 121.8 ± 24.4 , 80.5 ± 10.2 and $24.5 \pm 5.3 \mu\text{M}$, respectively (mean \pm s.e.m., Supplementary Fig. 6b, c). In contrast, repeat 1 and fragments AB and BC exhibited low nanomolar IC_{50} values towards m-calpain (Supplementary Fig. 6d). Our results confirm the specificity of calpastatin for the heterodimeric m- and μ -calpains and suggest that calpains that lack the small subunit, DIV and/or DIII may not support the same set of inhibitory interactions with calpastatin⁸.

The calpain and calpastatin proteins represent a major ubiquitous cellular proteolytic system, the imbalance of which has been implicated in necrosis associated with stroke and neuronal injury and perhaps Alzheimer's disease, heart disease, cataract formation, type 2 diabetes, cancer and LGMD-2A²⁹. Our study shows the mechanisms of activation by Ca^{2+} and inhibition by calpastatin of m- and μ -calpains (Fig. 4). Additional mechanisms of regulation for the calpain-calpastatin system include phosphorylation, membrane targeting and differential localization (Supplementary Fig. 13). The details of calpastatin specificity for the heterodimers may aid the design of new therapeutic agents.

METHODS SUMMARY

Repeat 1 of calpastatin and the m-calpain heterodimer, composed of an intact catalytic subunit and lacking the Gly-rich DV of the regulatory subunit, were expressed in *Escherichia coli* separately and were purified to homogeneity by fast-performance liquid chromatography (FPLC). Excess radio-labelled (for NMR) or unlabelled inhibitor was mixed with the protease in the presence of Ca^{2+} followed by further FPLC purification of the complex. The complex was subjected to limited proteolysis, NMR spectroscopy analysis, crystallization and X-ray crystallography structure determination. On the basis of the NMR assignments a shorter calpastatin fragment was generated and only its complex to m-calpain produced crystals. X-ray diffraction data collection was performed at three different US synchrotrons to improve the resolution. The crystal structure was obtained by molecular replacement using domains from previously determined calpain structures as search models. The active conformation of calpain and the mechanism of inhibition by calpastatin were probed by

site-directed mutagenesis using standard calpain activity and calpastatin inhibition assays on the basis of calpain autolysis and hydrolysis of fluorogenic substrates. Truncated calpastatin constructs and synthetic peptides were tested for inhibitory specificity against full-length and truncated calpain.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 6 March; accepted 15 August 2008.

1. Franco, S. J. & Huttenlocher, A. Regulating cell migration: calpains make the cut. *J. Cell Sci.* **118**, 3829–3838 (2005).
2. Wang, K. K. Calpain and caspase: can you tell the difference? *Trends Neurosci.* **23**, 20–26 (2000).
3. Harris, F., Biswas, S., Singh, J., Dennison, S. & Phoenix, D. A. Calpains and their multiple roles in diabetes mellitus. *Ann. NY Acad. Sci.* **1084**, 452–480 (2006).
4. Liu, J., Liu, M. C. & Wang, K. K. Calpain in the CNS: from synaptic function to neurotoxicity. *Sci. Signal.* **1**, re1 (2008).
5. Kramerova, I., Beckmann, J. S. & Spencer, M. J. Molecular and cellular basis of calpainopathy (limb girdle muscular dystrophy type 2A). *Biochim. Biophys. Acta* **1772**, 128–144 (2007).
6. Wendt, A., Thompson, V. F. & Goll, D. E. Interaction of calpastatin with calpain: a review. *Biol. Chem.* **385**, 465–472 (2004).
7. Moldoveanu, T. et al. A Ca^{2+} switch aligns the active site of calpain. *Cell* **108**, 649–660 (2002).
8. Suzuki, K., Hata, S., Kawabata, Y. & Sorimachi, H. Structure, activation, and biology of calpain. *Diabetes* **53**, S12–S18 (2004).
9. Hosfield, C. M., Elce, J. S., Davies, P. L. & Jia, Z. Crystal structure of calpain reveals the structural basis for Ca^{2+} -dependent protease activity and a novel mode of enzyme activation. *EMBO J.* **18**, 6880–6889 (1999).
10. Elce, J. S., Hegadorn, C., Gauthier, S., Vince, J. W. & Davies, P. L. Recombinant calpain II: improved expression systems and production of a C105A active-site mutant for crystallography. *Protein Eng.* **8**, 843–848 (1995).
11. Todd, B. et al. A structural model for the inhibition of calpain by calpastatin: crystal structures of the native domain VI of calpain and its complexes with calpastatin peptide and a small molecule inhibitor. *J. Mol. Biol.* **328**, 131–146 (2003).
12. Bode, W. & Huber, R. Structural basis of the endoprotease–protein inhibitor interaction. *Biochim. Biophys. Acta* **1477**, 241–252 (2000).
13. Cuerrier, D., Moldoveanu, T. & Davies, P. L. Determination of peptide substrate specificity for μ -calpain by a peptide library-based approach: the importance of primed side interactions. *J. Biol. Chem.* **280**, 40632–40641 (2005).
14. Cuerrier, D. et al. Development of calpain-specific inactivators by screening of positional scanning epoxide libraries. *J. Biol. Chem.* **282**, 9600–9611 (2007).
15. Moldoveanu, T., Campbell, R. L., Cuerrier, D. & Davies, P. L. Crystal structures of calpain–E64 and –leupeptin inhibitor complexes reveal mobile loops gating the active site. *J. Mol. Biol.* **343**, 1313–1326 (2004).
16. Betts, R., Weinsheimer, S., Blouse, G. E. & Anagali, J. Structural determinants of the calpain inhibitory activity of calpastatin peptide B27-WT. *J. Biol. Chem.* **278**, 7800–7809 (2003).
17. Betts, R. & Anagali, J. The beta- and gamma-CH2 of B27-WT's Leu11 and Ile18 side chains play a direct role in calpain inhibition. *Biochemistry* **43**, 2596–2604 (2004).
18. Moldoveanu, T., Hosfield, C. M., Jia, Z., Elce, J. S. & Davies, P. L. Ca^{2+} -induced structural changes in rat m-calpain revealed by partial proteolysis. *Biochim. Biophys. Acta* **1545**, 245–254 (2001).
19. Kiss, R. et al. Calcium-induced tripartite binding of intrinsically disordered calpastatin to its cognate enzyme, calpain. *FEBS Lett.* **582**, 2149–2154 (2008).

20. Reverter, D. *et al.* Flexibility analysis and structure comparison of two crystal forms of calcium-free human m-calpain. *Biol. Chem.* **383**, 1415–1422 (2002).
21. Strobl, S. *et al.* The crystal structure of calcium-free human m-calpain suggests an electrostatic switch mechanism for activation by calcium. *Proc. Natl Acad. Sci. USA* **97**, 588–592 (2000).
22. Pal, G. P., De Veyra, T., Elce, J. S. & Jia, Z. Crystal structure of a micro-like calpain reveals a partially activated conformation with low Ca^{2+} requirement. *Structure* **11**, 1521–1526 (2003).
23. Davis, T. L. *et al.* The crystal structures of human calpains 1 and 9 imply diverse mechanisms of action and auto-inhibition. *J. Mol. Biol.* **366**, 216–229 (2007).
24. Moldoveanu, T., Hosfield, C. M., Lim, D., Jia, Z. & Davies, P. L. Calpain silencing by a reversible intrinsic mechanism. *Nature Struct. Biol.* **10**, 371–378 (2003).
25. Blanchard, H. *et al.* Structure of a calpain Ca^{2+} -binding domain reveals a novel EF-hand and Ca^{2+} -induced conformational changes. *Nature Struct. Biol.* **4**, 532–538 (1997).
26. Lin, G. D. *et al.* Crystal structure of calcium bound domain VI of calpain at 1.9 Å resolution and its role in enzyme assembly, regulation, and inhibitor binding. *Nature Struct. Biol.* **4**, 539–547 (1997).
27. Dutt, P., Arthur, J. S., Grochulski, P., Cygler, M. & Elce, J. S. Roles of individual EF-hands in the activation of m-calpain by calcium. *Biochem. J.* **348**, 37–43 (2000).
28. Jia, Z. *et al.* Mutations in calpain 3 associated with limb girdle muscular dystrophy: analysis by molecular modeling and by mutation in m-calpain. *Biophys. J.* **80**, 2590–2596 (2001).
29. Saez, M. E., Ramirez-Lorca, R., Moron, F. J. & Ruiz, A. The therapeutic potential of the calpain family: new aspects. *Drug Discov. Today* **11**, 917–923 (2006).
30. Hanna, R. A., Campbell, R. L. & Davies, P. L. Calcium-bound structure of calpain and its mechanism of inhibition by calpastatin. *Nature* doi:10.1038/nature07451 (this issue).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank P. L. Davies for facilitating the initial stages of this work and for sharing the coordinates for the calpastatin repeat 4–calpain complex, M. Osborne for help with the NMR analysis, B. Schulman for collecting the ALS synchrotron data, the staff at Brookhaven National Light Source beam X29 and APS SERCAT for help with data collection, S. White and D. Miller for advice and sharing of synchrotron time, and C. Hosfield, A. Tocilj and R. Kriwacki for reviewing our manuscript. Funding was provided by the US National Institute of Health and National Cancer Institute, the Canadian Institute of Health Research and the American Lebanese Syrian Associated Charities.

Author Contributions T.M. performed the experimental work and data analysis. T.M., K.G. and D.R.G. were involved with the project planning and wrote the paper.

Author Information The sequences and three-dimensional coordinates for the calcium-dependent complex between m-calpain and calpastatin have been deposited in the PDB database under accession number 3DF0. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.R.G. (douglas.green@stjude.org).

METHODS

Cloning, mutagenesis, peptide synthesis, protein expression and purification.

The rat calpastatin repeat 1 clone encoding residues Met119–Ser238 (Supplementary Fig. 2b, gi 13540322) was provided by S. Arthur as an N-terminally His₁₀-tagged construct in pET16b vector. Subsequent cloning in the NcoI and XhoI sites of the kanamycin-resistant pET24d vector was performed using this vector as PCR template and the following oligonucleotides as primers, and engineering a stop codon to exclude the C-terminal His₆ tag: 5' med GCATGGCCATGGACAAGTCAGGCGTGAATGCTG, 3' med GTGGTGCTCGAGTTACTTTCCAGTTGGAGAGCTACAG, 5' sh GCATGGCCATGGCTGCTTTGGATGACCTGATAG, 3' sh GTGGTGCTCGAGTTAGGTGAAATCAGATGACCAGGCA, 5' BC GCATGGCCATGGACCAATGGATTCTACTAC and 3' BC GTGGTGCTCGAGTTAACAGGTGAAATCAGATGACAAGGC. We produced medium-sized (Met-Asp128–Lys226) and short (Met-Ala134–Thr219) calpastatin repeat 1 constructs and peptide BC (Met-Asp163–Cys220). In the short construct, we introduced a stop codon after Lys190 using the Quick Change protocol (Stratagene) and the forward primer GAACTTCTGGAGAAATAAGAAGCTATCACAGG (reverse not shown) to generate peptide AB (Met-Asp128–Lys190). The *E. coli* BL21 DE3 strain was used to express all five derivatives of calpastatin. In all calpastatin constructs the initiating Met was removed during expression as indicated by intact mass determination by mass spectrometry. The m-calpain heterodimer, C105S m80 kDa/21 kDa, which lacks the glycine-rich DV, and the protease core from μ -calpain (μ -II) were expressed in *E. coli* and purified as described previously^{7,10}. Forward mutagenesis primers for the 80 kDa subunit included R417A CCAGAAGCATCGGGCGCGCAGAGGAA, R420A GGCGGCGCAGGCGAAGATGGGTGAG and R469A CCTTCATCAACCTCGCGGAGGTCTCTAAC, and for calpastatin K176A GGCAGTGGGTATAGAAGGACTATTCC, E177A GCACTGGGTATAAAGGGACTATTCTC and K176/E177A GGCAGTGGGTATAGGGACTATTCTC. For ¹⁵N-labelling and ¹³C, ¹⁵N-labelling, the medium-sized and short calpastatins were expressed in M9 medium³¹ in the presence of ¹⁵N-NH₄Cl and ¹⁵N-NH₄Cl plus ¹³C-glucose, respectively. All forms of calpastatin were purified by boiling the cell lysate for 15 min, followed by Ni-NTA affinity chromatography (the original His-tagged construct), Q-Sepharose and C18 reversed-phase HPLC or S200 gel filtration chromatography. All protein preparations were exchanged into 20 mM HEPES (pH 7.5), 5 mM DTT storage buffer, were flash-frozen in liquid nitrogen, and stored at –80 °C. Peptide B1 Ac-ALGIKEGTIPPEYRKLL-NH₂ was synthesized and HPLC-purified at St Jude Children's Research Hospital Hartwell Center.

Calpain–calpastatin complex formation and purification. The m-calpain–calpastatin complex was formed by slowly titrating 50 mM CaCl₂ into a solution (~20–30 ml) containing purified calpain (10–20 mg) and a 2–5× molar excess of calpastatin in 50 mM HEPES buffer (pH 7.5) until the CaCl₂ concentration reached ~5–10 mM. The subsequent steps of purification were done in solutions containing 5 mM CaCl₂ to prevent dissociation of the complex. Excess untagged calpastatin was removed by recovering the complex on Ni-NTA (Qiagen), which bound to the column using the C-terminal His₆-tag on the calpain large subunit. Further purification involved Sephacryl S200 and Q-Sepharose chromatography. The complexes were stored in 20 mM HEPES (pH 7.5), 5 mM DTT, 10 mM CaCl₂ at –80 °C.

Limited proteolysis and autolysis of the calpain–calpastatin complex. Limited proteolysis¹⁸ or autolysis reactions were performed at 22 °C in a final volume of less than 150 μ l, in 1 mM CaCl₂ and 50 mM HEPES at pH 7.5. Purified 1 mg ml^{–1} (9 μ M) m-calpain–calpastatin complex was proteolysed by 0.001 mg ml^{–1} (0.04 μ M) trypsin (Sigma). For autolysis, 1 mg ml^{–1} (10 μ M) m-calpain was

incubated at 1 mM CaCl₂ with wild-type or mutant full-length calpastatin or fragments AB and BC. At specific time intervals, aliquots were removed, the reaction was stopped by the addition of 2× SDS gel sample buffer, and the products were analysed by SDS PAGE.

Calpastatin inhibition assays. Hydrolysis assays by m-calpain (1 nM or 20 nM), m-calpain mutants (1 nM) and μ -II (1.25–2.5 μ M) were performed in 100- μ l volumes containing 50 mM HEPES (pH 7.5), 200 mM NaCl, 1 mM DTT, various concentrations of substrate (0–1.5 mM SLY-MCA (Sigma) or 0–50 μ g ml^{–1} BODIPY-casein (Invitrogen)) and increasing inhibitor concentrations: 0–50 nM wild-type or mutant full-length calpastatin, fragments AB and BC, 0–20 μ M peptide B1 for m-calpain, and 0–750 μ M peptide B1 for μ -II. Duplicates or triplicates were performed for each condition in 96-well plates using a Molecular Devices microplate reader.

NMR assignments of calpastatin flexible regions in complex with calpain. The complexes between ¹⁵N-labelled medium or short calpastatin and unlabelled calpain were subjected to ¹⁵N, ¹H HSQC analysis in the storage buffer supplemented with 10% D₂O at 25 °C. The spectra were collected on a Bruker Avance DRX 600 MHz spectrometer equipped with a triple-resonance CryoProbe. HNCA and CBCA(CO)NH experiments with ¹⁵N/¹³C-labelled Asp128–Lys226 calpastatin bound to unlabelled calpain were used to assign the mobile regions of the inhibitor. Spectral processing was done using NMRpipe³² and spectral analysis using nmrView³³.

Calpain–calpastatin complex crystallization and structure determination.

After extensive screening and expansions, the complex between m-calpain and the NMR-trimmed Ala134–Thr219 calpastatin was crystallized in 4–9% PEG 3350, 5–10 mM CaCl₂ and 50–100 mM NaOAc (pH 5.5) using the hanging-drop method by mixing equal volumes of complex (10–15 mg ml^{–1}) and mother liquor. Cryo conditions included mother liquor and up to 30% ethylene glycol. The crystals were assigned to the tetragonal space group P₄2 with one molecule per asymmetric unit and diffracted to 2.8–3.5 Å at synchrotrons (Supplementary Table 1). Multiple native data sets were collected at 5.0.1–3 beamlines of the Advanced Light Source, X29 beamline at Brookhaven National Laboratories, and at SERCAT 22BM and 22ID beamlines at the Advanced Photon Source. The structure was determined by molecular replacement using as search models the Ca²⁺-bound rat m-calpain protease core (1MDW), the DIII and DIV of Ca²⁺-free rat m-calpain (1DF0) and the Ca²⁺-bound small subunit of rat m-calpain (1DVI) and Phaser³⁴ included in the CCP4 package³⁵. Structural refinement was processed using Refmac5 (ref. 36) and CNS³⁷, with manual fitting performed using Xfit³⁸.

31. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory Press, 1989).
32. Delaglio, F. *et al.* NMRPipe: a multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
33. Johnson, B. A. Using NMRView to visualize and analyze the NMR spectra of macromolecules. *Methods Mol. Biol.* **278**, 313–352 (2004).
34. McCoy, A. J. *et al.* Phaser crystallographic software. *J. Appl. Cryst.* **40**, 658–674 (2007).
35. Collaborative Computational Project, Number 4. The CCP4 suite: Programs for Protein Crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
36. Winn, M. D., Isupov, M. N. & Murshudov, G. N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. D* **57**, 122–133 (2001).
37. Brünger, A. T. *et al.* Crystallography and NMR system: a new software system for macromolecular structure determination. *Acta Crystallogr. D* **54**, 905–921 (1998).
38. McRee, D. E. A visual protein crystallographic software system for X11/XView. *J. Mol. Graph.* **10**, 44–46 (1992).

Calcium-bound structure of calpain and its mechanism of inhibition by calpastatin

Rachel A. Hanna¹, Robert L. Campbell¹ & Peter L. Davies¹

Calpains are non-lysosomal calcium-dependent cysteine proteinases that selectively cleave proteins in response to calcium signals¹ and thereby control cellular functions such as cytoskeletal remodeling, cell cycle progression, gene expression and apoptotic cell death^{2–4}. In mammals, the two best-characterized members of the calpain family, calpain 1 and calpain 2 (μ -calpain and m-calpain, respectively), are ubiquitously expressed. The activity of calpains is tightly controlled by the endogenous inhibitor calpastatin, which is an intrinsically unstructured protein capable of reversibly binding and inhibiting four molecules of calpain, but only in the presence of calcium^{5,6}. To date, the mechanism of inhibition by calpastatin and the basis for its absolute specificity have remained speculative^{7–9}. It was not clear how this unstructured protein inhibits calpains without being cleaved itself, nor was it known how calcium induced changes that facilitated the binding of calpastatin to calpain. Here we report the 2.4-Å-resolution crystal structure of the calcium-bound calpain 2 heterodimer bound by one of the four inhibitory domains of calpastatin. Calpastatin is seen to inhibit calpain by occupying both sides of the active site cleft. Although the inhibitor passes through the active site cleft it escapes cleavage in a novel manner by looping out and around the active site cysteine. The inhibitory domain of calpastatin recognizes multiple lower affinity sites present only in the calcium-bound form of the enzyme, resulting in an interaction that is tight, specific and calcium dependent. This crystal structure, and that of a related complex¹⁰, also reveal the conformational changes that calpain undergoes on binding calcium, which include opening of the active site cleft and movement of the domains relative to each other to produce a more compact enzyme.

The importance of the calpain–calpastatin balance in protection from pathological conditions where calpain is over-activated is well documented. Neurons could be protected from excitotoxic cell death after a glutamate-stimulated rise in intracellular calcium levels, either by overexpression of calpastatin or by expression of a sodium/calcium exchanger that was not cleavable by calpain¹¹. In rat hearts subjected to ischaemia/reperfusion injury, overexpression of calpastatin decreased troponin I degradation and contractile dysfunction¹². These studies benefit from the absolute specificity of calpastatin's inhibition of calpain, something lacking with the widespread use of low molecular weight calpain inhibitors that typically show some inhibition of other cysteine proteases¹³.

Calpains 1 and 2 are 100 kDa heterodimers with homologous large subunits comprising four domains (DI–DIV) and a common small subunit with two domains (DV, DVI) (Supplementary Fig. 1). The two carboxy-terminal penta-EF hand (PEF) domains, DIV and DVI, form the major heterodimer interface that holds the two subunits together through the pairing of their fifth EF-hands^{14,15}. Each of calpastatin's four inhibitory domains contains three regions of conserved sequence; subdomains A and C bind to DIV and DVI, respectively¹⁶,

and subdomain B inhibits calpain via a previously unknown mechanism. As the B subdomain was ineffective at inhibiting the isolated protease core¹⁷ and could be cross-linked to DIII¹⁸, it seemed that the inhibition of the whole enzyme might be allosteric. Others had speculated that the B subdomain might in some way block the active site cleft⁸.

To resolve these issues, we co-crystallized the inactive mutant of calpain 2 (C105S)¹⁹ with calpastatin inhibitory domain 4 (CAST4) in the presence of calcium (Supplementary Table 1). The 94-residue CAST4 binds as an extended polypeptide over the surface of calpain, making contact with each domain of the enzyme (Fig. 1, Supplementary Fig. 1) and burying approximately 2,800 Å² of the surface of calpain (Supplementary Table 2). The calcium-bound structure of calpain is more compact than the apo-form^{14,15} because the PEF domains are shifted towards the protease core (DI, DII), decreasing the radius of gyration from 32.2 Å to 30.8 Å (Supplementary Fig. 2). In the process, the amino-terminal anchor helix is displaced from its binding site on the small subunit near EF-hand 1 and is not seen in the electron density, consistent with its tendency to be autoproteolysed during activation of the enzyme.

When bound to calpain, subdomain C forms an amphipathic α -helix that binds to an exposed hydrophobic patch on DVI, just as

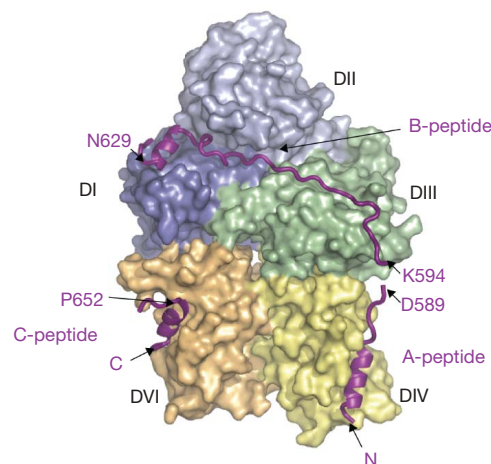


Figure 1 | Overview of calpastatin domain 4 (CAST4) bound to calpain 2. The overall structure of CAST4 (purple) bound to the inactive C105S mutant of calpain 2. CAST4, which is unstructured in the absence of calpain, forms three α -helices when in complex with the enzyme. Helices in subdomains A and C, which are in contact with DIV (yellow) and DVI (orange), and the helix in subdomain B, which is in contact with the protease core DI and DII (blue and light blue, respectively) are shown in ribbon representation. DIII is coloured green. Gaps in the electron density of CAST4 are indicated by missing residues between D589 and K594, and between N629 and P652.

¹Department of Biochemistry, Queen's University, Kingston, Ontario, Canada K7L 3N6.

previously described⁸, and subdomain A forms a similar amphipathic α -helix that interacts with DIV on the large subunit (Supplementary Fig. 3, Supplementary Tables 3 and 4). It is the B subdomain of calpastatin that inhibits calpain²⁰ while subdomains A and C increase the overall affinity of the interaction. From the crystal structure we see that the B subdomain binds to the activated enzyme on either side of the active site cleft (Fig. 2a, Supplementary Fig. 4, Supplementary Table 5). On the N-terminal side, the inhibitor forms hydrophobic and electrostatic interactions with a shallow groove in DIII (Fig. 2a) that becomes aligned with the catalytic cleft in the holo form of the enzyme. Continuing on into the unprimed side of the cleft, the highly conserved Leu 612 of CAST4 occupies the hydrophobic pocket of the S2 subsite (Fig. 2b, 3b). This leucine is strongly favoured in m-calpain substrates²¹, and frequently appears in this position in inhibitors—for example, leupeptin (Fig. 3d). Immediately after Leu 612 the backbone of Gly 613 twists sharply away from the active site and traces the path of what would normally be the side chain of the P1 residue (Fig. 3a, b, Supplementary Fig. 5a, b). In this way, the potentially scissile bond between the P1 and P1' residues is held ~ 2 Å out of reach of calpain's catalytic *Cys 105 (asterisk indicates calpain residues) (Fig. 3c). This is demonstrated in our structure, by the proximity of the bond to the *Ser 105 replacement for *Cys 105, which, because this is a calcium-bound structure, is correctly aligned with *His 262 and *Asn 286 of the triad for catalysis.

On the C-terminal side of the B subdomain, calpastatin forms a two-turn amphipathic α -helix that binds to DI (Fig. 2d). Another feature of the C-terminal side of the B subdomain is a seven-residue highly conserved TIPPEYR sequence. The two prolines, Pro 620 and Pro 621, break the helix and redirect CAST4 into the primed side of the cleft while making key contacts with *Trp 288 in DII (Fig. 2c, Supplementary Fig. 6). *Trp 288 is an important residue in the calcium-dependent activation of the protease core, as it reorients from lying across the cleft as a wedge to lining the pocket of the cleft¹⁷. In particular, the conserved Pro 620 stacks against *Trp 288 and helps pin the C-terminal side of the loop on the primed side of the cleft. Additional hydrophobic contacts come from Ile 619 which projects

into the cleft and contacts *Leu 102 and *Ala 101 (Supplementary Fig. 6). Thr 618 (in the TIPPEYR sequence) mirrors Gly 613 by exiting the cleft, oriented such that its side chain projects where the main chain should be, and vice versa (Fig. 2c).

The Leu 612–Gly 613 dipeptide is well conserved across the four inhibitory domains of calpastatin and across a variety of species (Gly 613 is absolutely conserved), as is the TIPPEYR sequence (Supplementary Fig. 7). In between these conserved sequences is a loop out of four typically hydrophilic residues (ERDD) between G and T that are not as well conserved. The KLGE sequence (residues 611–614) forms a type-II β -turn as suggested in ref. 22. The RDDTI sequence (residues 615–619) forms two overlapping, distorted type-I β -turns. For this region to loop out, the flanking portions of the B subdomain (coloured purple and green in Fig. 2a) must each tightly and independently bind to calpain.

This is indeed a novel mechanism. Among proteinaceous inhibitors of proteases, calpastatin is most similar to cystatins that act as a wedge to block the active site and inhibit cathepsins²³. Cystatins bind as two hairpin loops to both the primed and unprimed side of the active site cleft. Calpastatin is one of only two protease inhibitors that are known to be intrinsically unstructured proteins when not bound to their cognate targets. The other is IA3, an inhibitor of yeast proteinase A, but unlike calpastatin, IA3 is a short (68-residue) polypeptide that is induced to form a nearly perfect helix that blocks the active site of its target²⁴. However, in contrast to all other protein inhibitors, calpastatin binds as a contiguous stretch of polypeptide across the complete cleft in such a way that the catalytic cysteine is prevented from reaching the P1 carbonyl. A similar situation exists with the protein phosphatase 1–inhibitor 2 (PP1/I-2) complex²⁵. Like calpastatin, I-2 is an intrinsically disordered protein that becomes ordered on binding. It also contacts the enzyme in three regions, but is bound in an almost entirely α -helical conformation, unlike calpastatin which maintains a large segment containing no regular secondary structure.

To test our model for the inhibition of calpain by calpastatin, we made three mutations. In the first of these the conserved Gly 613 was

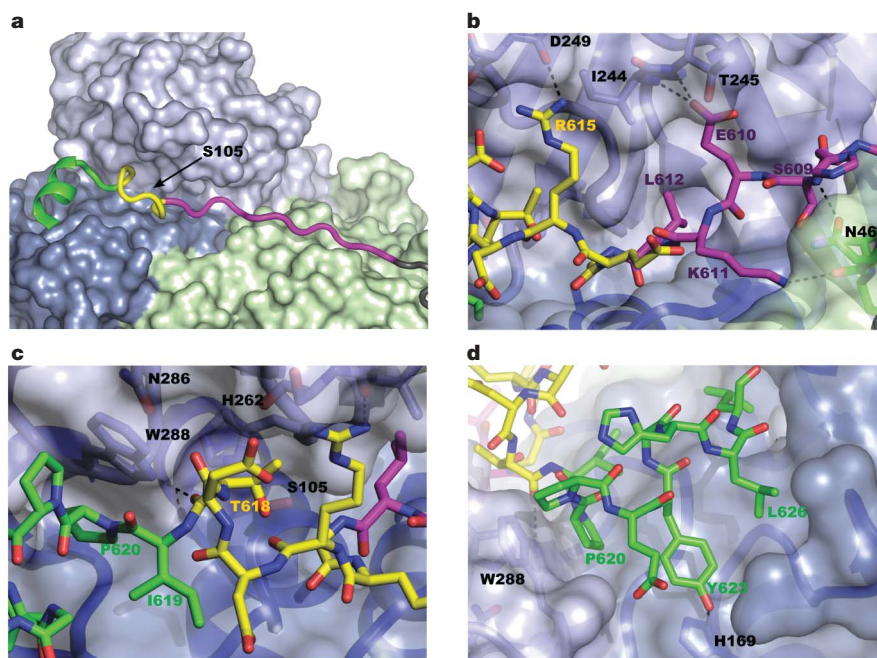


Figure 2 | Specific interactions of calpastatin with calpain entering and leaving the active-site cleft. **a–d**, The 27-residue B-peptide⁷ is coloured as follows: the residues that make the loop out of the active site are coloured yellow, the residues N-terminal to the loop are purple, and the residues C-terminal to the loop are green. Other calpastatin residues are coloured dark grey. Hydrogen-bond interactions of calpastatin with calpain (coloured

as in Fig. 1) are shown by black dashed lines. O and N atoms are coloured red and blue, respectively. **a**, Overview of calpain binding at the active site of calpain. **b**, Close-up view of the calpastatin at the unprimed side of the active site. **c**, Close-up view of calpastatin looping away from the catalytic residue. **d**, Close-up view of calpastatin at the primed side of the active site.

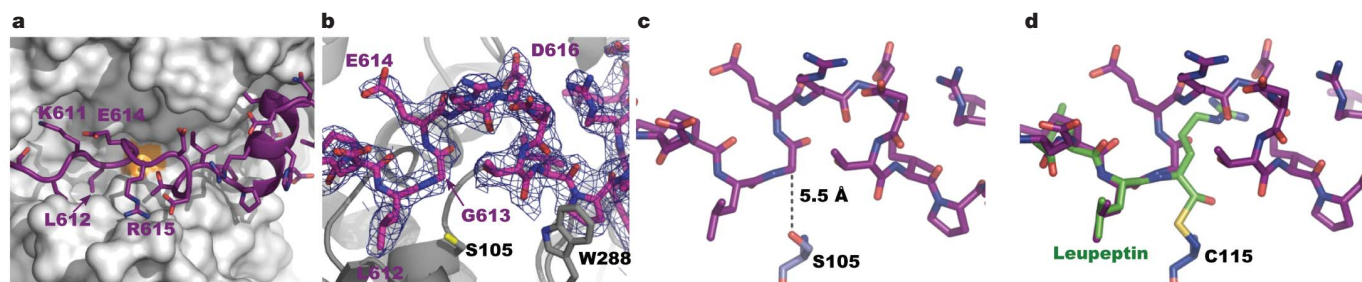


Figure 3 | Close-up views of calpastatin at the active site. **a**, Representation of calpastatin (purple) looping away from the mutant active-site serine (orange) of calpain (shown uncoloured in space-filling mode). Calpastatin side chains are shown with O atoms in red and N atoms in blue. **b**, Electron density at 1σ for a $2F_o - F_c$ map of the loop region (blue mesh). Calpastatin (purple) backbone O and N atoms are also coloured red and blue, respectively. Calpain (grey) side chains are identified with black lettering. The side chain O atom of *S105 is coloured yellow. **c**, **d**, Comparison of the

calpastatin loop region (**c**) with the superimposed leupeptin-bound structure of calpain 1 protease core²⁸ (**d**). The distance between the active-site serine and the P1 carbonyl of calpastatin is shown by the dotted line (**c**). The structure of leupeptin (green) covalently linked to *Cys115 of calpain 1 is superimposed on the calpastatin loop (**d**) to show how closely the backbone at Gly 613 follows the side chain of the P1 arginine of leupeptin. The P2 leucine of both CAST4 and leupeptin overlap closely.

mutated to Ala to reduce backbone flexibility, rendering CAST4 less able to loop out of the active-site cleft. This single amino acid substitution increased the IC_{50} value 2.9-fold, from 64 to 182 nM (Supplementary Figs 8, 9a). Second, we inserted a Phe residue between Leu 612 and Gly 613 (insert-F) to prevent the formation of the type II β -turn in CAST4 that holds the glycine away from the catalytic cysteine while placing phenylalanine in the P1 position where its peptide bond with glycine should be cleavable. This mutation turned CAST4 from an inhibitor into a substrate. Within a minute of forming the complex, the mutant CAST4 was cut into two fragments of mass 5,068.6 and 5,745.1 Da, which correspond to residues 570–613 and 614–664, respectively (Supplementary Fig. 9b). Interestingly, despite being rapidly cleaved this mutant is still a potent inhibitor of calpain with an IC_{50} of 112 nM, approximately 1.8 times larger than that of the wild type, which makes it a better inhibitor than the G613A mutant (Supplementary Fig. 8). A third mutation, *A101D, was designed to spoil a contact calpastatin makes with the protease core. In calpain 2, *Ala 101 is in a hydrophobic patch that binds calpastatin after the conserved TIPPEYR sequence. *Ala 101 was mutated to aspartic acid to test that we had identified the binding site of subdomain B on calpain. The *A101D mutation caused a 2.2-fold increase in the IC_{50} value of wild-type CAST4 to 146 nM (Supplementary Fig. 8).

In addition to uncovering this mechanism of protease inhibition, the calpastatin–calpain structure reveals for the first time the calcium-bound structure of calpain (Supplementary Fig. 2). The domains have changed little from the calcium-free form. What have changed in the calcium-bound structure of calpain are the relative positions of the domains. In particular, the two PEF domains are closer to the core and have displaced the N-terminal anchor helix. It is not clear to what extent release of the anchor helix helps bring the PEF domains closer to the core and to what extent the domain movement displaces the helix. The overall structural changes are illustrated by a simulation in which Ca^{2+} ions are docked to their binding sites (Supplementary Movie 1). At the 5 mM $CaCl_2$ concentration used for crystallization, four calcium ions are present in both PEF domains. Those in DVI of the small subunit are in exactly the same locations they were in the calcium-bound DVI homodimer structure⁸. Those in DIV of the large subunit are in equivalent locations. Rather surprisingly, despite reports that isolated recombinant DIII can bind Ca^{2+} , there were no calcium ions in contact with this domain in the CAST4-bound structure²⁶. The protease core contains the two calcium ions, one in each domain, that are responsible for the cooperative assembly of the catalytic cleft²⁷. The arrangement of the core within the whole enzyme is exactly that seen in the crystal structure of the calcium-bound core alone, which validates the mechanism of activation that was elucidated with mutants of the isolated core²⁷.

By comparing the calpain–calpastatin structure to the calcium-bound structures of the protease core and PEF DVI we see indications that calpastatin recognizes the calcium-bound structure, rather than causing structural changes on binding. As mentioned, subdomains A and C bind when calcium causes a shift in the EF hands, exposing the hydrophobic areas. Without this shift there is not sufficient room for these subdomains to bind⁸. In our structure, the conserved hydrophobic residues of subdomains A and C occupy deep hydrophobic pockets in calpain (Supplementary Fig. 3b, c). Likewise, the structural changes that occur when the protease core binds calcium align the active site, and in particular causes a large change in the position of *Trp 288, without which the B sub-domain could not occupy and bind to the cleft. Also, the rearrangement of the domains relative to one another allows the simultaneous interaction of calpastatin with both DIII and the protease core. This increases the area of interaction, the number of low affinity sites along the length of CAST4, and thus the overall affinity of CAST for calpain. The fact that calpastatin binds only after calpain is activated raises the issue of how closely spaced these proteins are in the cell. Their proximity will probably control the window of opportunity that calpain has for proteolysis during calcium signalling before being inhibited.

The observed shift in DIII is also interesting from the standpoint of designing specific inhibitors to counter the overactivation of calpain that occurs in various diseases¹. Previous attempts at docking inhibitors have been forced to use the protease core²⁸, but the structure reported here will enable the design of inhibitors that can take advantage of the complete active site cleft.

METHODS SUMMARY

Inactive calpain 2 with its truncated small subunit (*C105S-m80k/21k) was expressed and purified as previously described¹⁹. Wild-type CAST4 and the mutants were produced in *Escherichia coli* BL21 (DE3) under kanamycin selection and purified using heat denaturation and HPLC.

To produce the calpain–CAST4 complex, 1 M $CaCl_2$ was added at $6\ \mu\text{l min}^{-1}$ to a mixture of calpain (~15 mg) and a twofold molar excess of CAST4 in 50 ml. The complex was purified using two column chromatography steps: a Ni^{2+} -chelating column (Qiagen), and Superdex 200 C 26/100 column (GE Healthcare). The pure complex was concentrated to $30\text{--}40\ \text{mg ml}^{-1}$ using an Amicon Ultra-15 30k molecular weight cut-off centrifugal filter (Millipore), aliquoted and flash frozen. The protein solution was added to an equal volume of precipitant solution of 100 mM MES (pH 6.5), 25% PEG 1000 (Qiagen).

Crystallization of the calpain 2 *C105S-m80k/21k–CAST4 complex was performed by the microbatch method under paraffin oil (Hampton). Diffraction data on the calpain–CAST4 complex were collected on a single crystal with approximate dimensions $10 \times 50 \times 100\ \mu\text{m}$ at the X6A beamline of the NSLS facility (Brookhaven National Laboratory). The structure was determined by molecular replacement using the structure of the calcium-bound protease core¹⁷ and the domains DIII, DIV and DVI of the structure of calcium-free full-length calpain¹⁵.

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 11 March; accepted 24 September 2008.

- Goll, D. E., Thompson, V. F., Li, H., Wei, W. & Cong, J. The calpain system. *Physiol. Rev.* **83**, 731–801 (2003).
- Glading, A., Lauffenburger, D. A. & Wells, A. Cutting to the chase: Calpain proteases in cell motility. *Trends Cell Biol.* **12**, 46–54 (2002).
- Franco, S. J. & Huttenlocher, A. Regulating cell migration: Calpains make the cut. *J. Cell Sci.* **118**, 3829–3838 (2005).
- Abe, K. & Takeichi, M. NMDA-receptor activation induces calpain-mediated beta-catenin cleavages for triggering gene expression. *Neuron* **53**, 387–397 (2007).
- Wendt, A., Thompson, V. F. & Goll, D. E. Interaction of calpastatin with calpain: A review. *Biol. Chem.* **385**, 465–472 (2004).
- Hanna, R. A., Garcia-Diaz, B. E. & Davies, P. L. Calpastatin simultaneously binds four calpains with different kinetic constants. *FEBS Lett.* **581**, 2894–2898 (2007).
- Betts, R., Weinsheimer, S., Blouse, G. E. & Anagli, J. Structural determinants of the calpain inhibitory activity of calpastatin peptide B27-WT. *J. Biol. Chem.* **278**, 7800–7809 (2003).
- Todd, B. *et al.* A structural model for the inhibition of calpain by calpastatin: Crystal structures of the native domain VI of calpain and its complexes with calpastatin peptide and a small molecule inhibitor. *J. Mol. Biol.* **328**, 131–146 (2003).
- Pfizer, J., Assfalg-Machleidt, I., Machleidt, W. & Schaschke, N. Inhibition of human mu-calpain by conformationally constrained calpastatin peptides. *Biol. Chem.* **389**, 83–90 (2008).
- Moldoveanu, T., Gehring, K. & Green, D. R. Concerted multi-pronged attack by calpastatin to occlude the catalytic cleft of heterodimeric calpains. *Nature* doi:10.1038/nature07353 (this issue).
- Bano, D. *et al.* Cleavage of the plasma membrane $\text{Na}^+/\text{Ca}^{2+}$ exchanger in excitotoxicity. *Cell* **120**, 275–285 (2005).
- Maekawa, A. *et al.* Overexpression of calpastatin by gene transfer prevents troponin I degradation and ameliorates contractile dysfunction in rat hearts subjected to ischemia/reperfusion. *J. Mol. Cell. Cardiol.* **35**, 1277–1284 (2003).
- Neffe, A. T. & Abell, A. D. Developments in the design and synthesis of calpain inhibitors. *Curr. Opin. Drug Discov. Dev.* **8**, 684–700 (2005).
- Hosfield, C. M., Elce, J. S., Davies, P. L. & Jia, Z. Crystal structure of calpain reveals the structural basis for Ca^{2+} -dependent protease activity and a novel mode of enzyme activation. *EMBO J.* **18**, 6880–6889 (1999).
- Strobl, S. *et al.* The crystal structure of calcium-free human m-calpain suggests an electrostatic switch mechanism for activation by calcium. *Proc. Natl Acad. Sci. USA* **97**, 588–592 (2000).
- Takano, E., Ma, H., Yang, H. Q., Maki, M. & Hatanaka, M. Preference of calcium-dependent interactions between calmodulin-like domains of calpain and calpastatin subdomains. *FEBS Lett.* **362**, 93–97 (1995).
- Moldoveanu, T. *et al.* A Ca^{2+} switch aligns the active site of calpain. *Cell* **108**, 649–660 (2002).
- Croall, D. E. & McGrody, K. S. Domain structure of calpain: Mapping the binding site for calpastatin. *Biochemistry* **33**, 13223–13230 (1994).
- Elce, J. S., Hegadorn, C., Gauthier, S., Vince, J. W. & Davies, P. L. Recombinant calpain II: Improved expression systems and production of a C105A active-site mutant for crystallography. *Protein Eng.* **8**, 843–848 (1995).
- Maki, M. *et al.* Inhibition of calpain by a synthetic oligopeptide corresponding to an exon of the human calpastatin gene. *J. Biol. Chem.* **264**, 18866–18869 (1989).
- Cuerrier, D., Moldoveanu, T. & Davies, P. L. Determination of peptide substrate specificity for mu-calpain by a peptide library-based approach: The importance of primed side interactions. *J. Biol. Chem.* **280**, 40632–40641 (2005).
- Ishima, R. *et al.* Structure of the active 27-residue fragment of human calpastatin. *FEBS Lett.* **294**, 64–66 (1991).
- Otlewski, J., Filip, J., Zakrzewska, M. & Oleksy, A. The many faces of protease-protein inhibitor interaction. *EMBO J.* **24**, 1304–1310 (2005).
- Li, M. *et al.* The aspartic proteinase from *Saccharomyces cerevisiae* folds its own inhibitor into a helix. *Nature Struct. Biol.* **7**, 113–117 (2000).
- Hurley, T. D. *et al.* Structural basis for regulation of protein phosphatase 1 by inhibitor-2. *J. Biol. Chem.* **282**, 28874–28883 (2007).
- Tompa, P., Emori, Y., Sorimachi, H., Suzuki, K. & Friedrich, P. Domain III of calpain is a Ca^{2+} -regulated phospholipid-binding domain. *Biochem. Biophys. Res. Commun.* **280**, 1333–1339 (2001).
- Moldoveanu, T., Jia, Z. & Davies, P. L. Calpain activation by cooperative Ca^{2+} binding at two non-EF-hand sites. *J. Biol. Chem.* **279**, 6106–6114 (2004).
- Moldoveanu, T., Campbell, R. L., Cuerrier, D. & Davies, P. L. Crystal structures of calpain-E64 and -leupeptin inhibitor complexes reveal mobile loops gating the active site. *J. Mol. Biol.* **343**, 1313–1326 (2004).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements This work was funded by the Canadian Institutes for Health Research. R.A.H. is the recipient of a Bauman Fellowship and an R.J. Wilson Fellowship. P.L.D. holds a Canada Research Chair in Protein Engineering. We are grateful to S. Gauthier for technical assistance, to M. Kuiper for the simulation of the conformational change in calpain on binding Ca^{2+} , to the Queen's University Protein Function Discovery (PFD) Facility and D. McLeod for mass spectrometry, Z. Jia for access to data collection on a home source and to M. Allaire at Brookhaven National Laboratory for assistance with data collection on beam line X6A.

Author Information Structure factors and coordinates have been deposited in the RCSB Protein Data Bank with the accession number 3BOW. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to P.L.D. (peter.davies@queensu.ca).

METHODS

Data analysis and structure solution. The oscillation data were processed using HKL2000²⁹ and the CCP4 software suite³⁰. The structure was determined by molecular replacement using the automated search method of phaser^{31,32}. The structure was refined with refmac5 using the TLS protocol³³.

Kinetic analysis. Calpain assays of the hydrolysis of a FRET-linked peptide substrate were performed in triplicate in a final volume of 1 ml. The concentrations of inhibitors ranged from 24 nM to 2.2 μ M. The concentration of wild-type and mutant calpain 2 was 50 nM. The end-point fluorescence intensity was measured in a LS 50B fluorimeter (Perkin Elmer). The IC₅₀ values were obtained by three-parameter sigmoidal fitting using the data graphing software SigmaPlot (Systat Software Inc.). Calpain was mixed with varying concentrations of CAST in 500 μ l of 50 mM HEPES-HCl (pH 7.4). Reactions were initiated by the addition of 500 μ l of 50 mM HEPES-HCl (pH 7.4), 2 mM CaCl₂, 2 mM DTT, and 10 μ M (EDANS)-EPLFAERK-(DABCYL), the fluorogenic substrate of calpain²¹. The reaction was allowed to proceed for 20 min and was stopped with 10 μ l of 0.5 M EDTA. Fluorescence intensity was measured using wavelengths of 335 nm for excitation and 505 nm for emission.

Limited proteolysis and identification of cleavage products of the insert-F mutant. Proteolytic digestion of the CAST4 insert-F mutant (40 nM) by calpain 2 (400 nM) was performed at 22 °C in 10 mM HEPES-HCl (pH 7.4), 1 mM dithiothreitol in 200 μ l total volume. Digestion was initiated by the addition of CaCl₂ to a concentration of 1 mM and was stopped at 1 min by the addition of EDTA to a concentration of 100 mM. Digestion of the insert-F mutant was accessed by 10% SDS-PAGE and Coomassie blue staining. The insert-F fragments were excised from the gel, and digestion analysis by MALDI-TOF mass spectrometry was performed at the Protein Function Discovery facility (Queen's University, Kingston, Canada).

29. Minor, W., Cymborowski, M., Otwinowski, Z. & Chruszcz, M. HKL-3000: The integration of data reduction and structure solution — from diffraction images to an initial model in minutes. *Acta Crystallogr. D* **62**, 859–866 (2006).
30. Collaborative Computational Project Number 4. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
31. Read, R. J. Pushing the boundaries of molecular replacement with maximum likelihood. *Acta Crystallogr. D* **57**, 1373–1382 (2001).
32. Storoni, L. C., McCoy, A. J. & Read, R. J. Likelihood-enhanced fast rotation functions. *Acta Crystallogr. D* **60**, 432–438 (2004).
33. Winn, M., Isupov, M. & Murshudov, G. N. Use of TLS parameters to model anisotropic displacements in macromolecular refinement. *Acta Crystallogr. D* **57**, 122–133 (2001).

The ion pathway through the opened Na^+, K^+ -ATPase pump

Ayako Takeuchi¹, Nicolás Reyes¹, Pablo Artigas¹ & David C. Gadsby¹

P-type ATPases pump ions across membranes, generating steep electrochemical gradients that are essential for the function of all cells. Access to the ion-binding sites within the pumps alternates between the two sides of the membrane¹ to avoid the dissipation of the gradients that would occur during simultaneous access. In Na^+, K^+ -ATPase pumps treated with the marine agent palytoxin, this strict alternation is disrupted and binding sites are sometimes simultaneously accessible from both sides of the membrane, transforming the pumps into ion channels (see, for example, refs 2, 3). Current recordings in these channels can monitor accessibility of introduced cysteine residues to water-soluble sulphhydryl-specific reagents⁴. We found previously⁵ that Na^+, K^+ pump-channels open to the extracellular surface through a deep and wide vestibule that emanates from a narrower pathway between transmembrane helices 4 and 6 (TM4 and TM6). Here we report that cysteine scans from TM1 to TM6 reveal a single unbroken cation pathway that traverses palytoxin-bound Na^+, K^+ pump-channels from one side of the membrane to the other. This pathway comprises residues from TM1, TM2, TM4 and TM6, passes through ion-binding site II, and is probably conserved in structurally and evolutionarily related P-type pumps, such as sarcoplasmic- and endoplasmic-reticulum Ca^{2+} -ATPases and H^+, K^+ -ATPases.

The Na^+, K^+ -ATPase is a P-type (named for its phosphorylated intermediate) pump that exports three Na^+ ions and imports two K^+ ions per ATP hydrolysed. The ion-binding sites are accessible from the extracellular space in the phosphorylated conformation, called E2P, and from the cytoplasm in the dephosphorylated configuration, E1. But the routes by which ions approach and leave those sites are still not understood⁶ despite the availability of X-ray crystal structures of sarcoplasmic- and endoplasmic-reticulum Ca^{2+} -ATPase (SERCA) P-type pumps in several states^{6–11}, although a recent structure of a BeF_3^- -trapped E2P-like state captured an open luminal pathway^{12,13}. However, sensitive electric current recording methods developed for studies of ion channels¹⁴ have begun to probe the ion pathway of the Na^+, K^+ pump^{5,15–18}, after its transformation into an ion channel by palytoxin² and electrophysiological analyses of reactivity of introduced cysteines to methanethiosulphonate (MTS) reagents⁴.

Cysteines modified by MTS reagents in palytoxin-bound Na^+, K^+ pump-channels include those substituted for ion-binding^{19–21} acidic residues in the pocket between TM4 (for example E336, equivalent to SERCA E309) and TM6 (for example D813, equivalent to SERCA N796), as well as T806 (equivalent to SERCA P789) at the outermost end of TM6 within the external vestibule floor^{5,16,17}. These three positions approximately align (Fig. 1) in extracellular views of the transmembrane domain of the Na^+, K^+ -ATPase, whether of the recent²¹ E2· MgF_4^{2-} Na^+, K^+ -ATPase structure (Fig. 1a), which is an occluded conformation with both cytoplasmic and extracellular pathways shut, or of a model based on the E2P-like SERCA E2· BeF_3^-

structure^{12,13} (Fig. 1b), in which the cytoplasmic pathway is shut but the extra-cytoplasmic pathway is open. However, palytoxin, with Na^+ and ATP present, appears to stabilize an E2P-related Na^+, K^+ pump conformation^{22,23} in which the gates to the binding sites can both be open³, a structure not yet visualized (or expected) for any native P-type pump. In occluded structures of SERCA containing two Ca^{2+} ions^{7,10,11} and of Na^+, K^+ -ATPase²¹ containing two K^+ ions, side chains of residues in TM5 and TM6 help coordinate the bound ion in site I, and TM4 and TM6 side chains help coordinate that in site II. The near alignment of accessible TM4 and TM6 positions (Fig. 1) therefore raises two questions: do ions in the Na^+, K^+ pump's extracellular pathway flow between TM4, TM6 and TM5 (ref. 17) or between TM4, TM6, TM2 and TM1 (compare with refs 6, 12, 13) (Fig. 1, red question marks); and what pathway(s) do ions take from the binding sites to the cytoplasm?

To answer these questions, we first introduced cysteines, one at a time, at 20 contiguous positions (1778–1797) along TM5 and at 4 more (A798–P801) in the external loop connecting TM5 and TM6, into the *Xenopus* Na^+, K^+ -ATPase α_1 subunit made ouabain resistant by the mutation²⁴ C113Y. We co-expressed each cysteine-tagged mutant with the *Xenopus* β_3 subunit in *Xenopus* oocytes^{5,18}. After applying 50 nM palytoxin (Fig. 2a–d, black arrowheads) to outside-out membrane patches, to transform all Na^+, K^+ pumps into ion channels (signals from native ouabain-sensitive *Xenopus* Na^+, K^+ pumps were prevented using 100 μM ouabain in all external solutions), we assessed

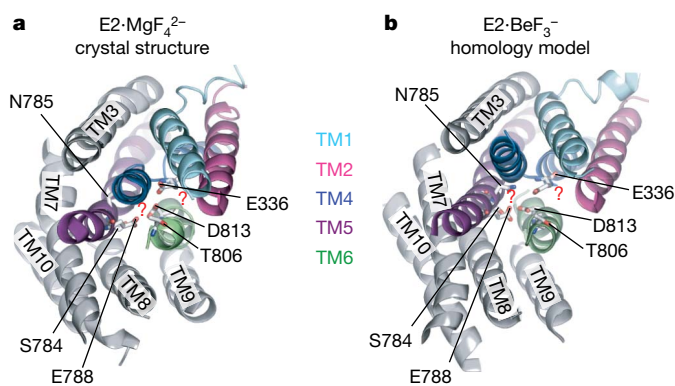


Figure 1 | Alternative routes for ions through the Na^+, K^+ -ATPase transmembrane domain. Extracellular views of the ten transmembrane helices of the Na^+, K^+ -ATPase E2· MgF_4^{2-} crystal structure²¹ (a; Protein Data Bank code 3B8E) and a homology model of the Na^+, K^+ -ATPase based on the SERCA E2· BeF_3^- structure¹² (b; Protein Data Bank code 3B9B). Helices are coloured grey except TM1 (pale blue), TM2 (magenta), TM4 (blue), TM5 (purple) and TM6 (green). Red question marks label two possible ion pathways: one between TM5, TM4 and TM6, and the other between TM4, TM1, TM2 and TM6. Key residues in these pathways are labelled.

¹Laboratory of Cardiac/Membrane Physiology, The Rockefeller University, New York, New York 10065, USA.

reactivity of positively charged, membrane-impermeant, 1 mM 2-trimethylammonium-ethyl-methanethiosulphonate (MTSET^+ ; Fig. 2a–d, blue arrows) with each engineered cysteine. Reaction was signalled by alteration of the inward Na^+ current (symmetrical 125 mM Na^+ solutions with -50-mV membrane potential) flowing through pump-channels (Fig. 2). MTSET^+ tests were preceded by exposure to 10 mM dithiothreitol (grey arrows) to restore any spontaneously oxidized thiols. There was no evidence of MTSET^+ reaction with any residue in TM5, but it rapidly decreased current by $\sim 25\%$ in construct N799C with a cysteine in the TM5–TM6 loop (Fig. 2a).

We similarly scanned 21 contiguous positions (F99–I119) in TM1, 11 (Y133–V142, T145) in TM2 and 6 (Q120, Q128–L132) in the extracellular TM1–TM2 connecting loop, testing reactivity of each introduced cysteine with 1 mM MTSET^+ (Fig. 3). Reactive positions (defined as $>10\%$ change in pump-channel current) in TM1 were G100, G101, F102, S103, L106, C113, A116 and Y117 (Fig. 3a–c), those in the TM1–TM2 loop included Q120, Q128, D130 and N131 (Fig. 3c, f), and those in TM2 included Y133 and L134 at the outer end and T145 towards the cytoplasmic end (Fig. 3d–f).

The summarized results from these scans, mapped onto an Na^+, K^+ pump homology model based on the SERCA $\text{E2} \cdot \text{BeF}_3^-$ structure, show, as red sticks, residues in positions where substituted cysteines showed evidence of modification by 1 mM MTSET^+ , and, as yellow sticks, residues in positions where there was no evidence of reactivity (Fig. 4a, b). Our previously reported⁵ results on cysteines introduced in TM4 (E321, E336, G337), the TM5–TM6 loop (L802–L804) and TM6 (G805–C811, D813, D817) are included. To fill gaps, we tested 16 additional strategically located positions in TM3 (I290, I294, I297, A301) and TM4 (A322–F325, I327, G328, V331, A332,

P335, L339, T341, V342): only P335C and L339C mutants showed reactivity with MTSET^+ .

The red residues mark out a single, unbroken MTSET^+ -accessible pathway (Fig. 4a, b; Supplementary Figs 1, 2) that runs between TM1, TM2, TM6 and TM4 rather than between TM5, TM4 and TM6 (Fig. 4a; Supplementary Figs 1, 2), passes through site II, and spans the full distance across the membrane (approximate boundaries indicated by lines $\sim 35 \text{ \AA}$ apart in Fig. 4b; see also Supplementary Fig. 2 and Supplementary Movie). Red reactive positions are enveloped in a yellow non-responsive surround (Fig. 4a, b; Supplementary Fig. 2a, b), indicating that the scan was complete and thus fully delimits this principal pathway through the pump. Moreover, as current was practically abolished after MTSET^+ modification of cysteine substitutes at TM1 position L106 (Figs 3a, c, 4c; Supplementary Figs 4a, 7c, g), or position G337 in TM4 (ref. 5) or T806 in TM6 (ref. 5; Supplementary Fig. 3b), the pathway depicted in Fig. 4 (and Supplementary Fig. 2) is probably the sole route for rapid ($\sim 10^7 \text{ s}^{-1}$) Na^+ ion flow through palytoxin-bound Na^+, K^+ pump-channels.

The negative charges of site-II residues E336 (TM4) and D813 (TM6), which are largely conserved in P-type cation pumps, form a cation-selectivity filter⁵. This was proposed to be responsible for the apparent lack of reactivity of a cysteine substituted for nearby G337 with negatively charged 2-sulphonato-ethyl-methanethiosulphonate

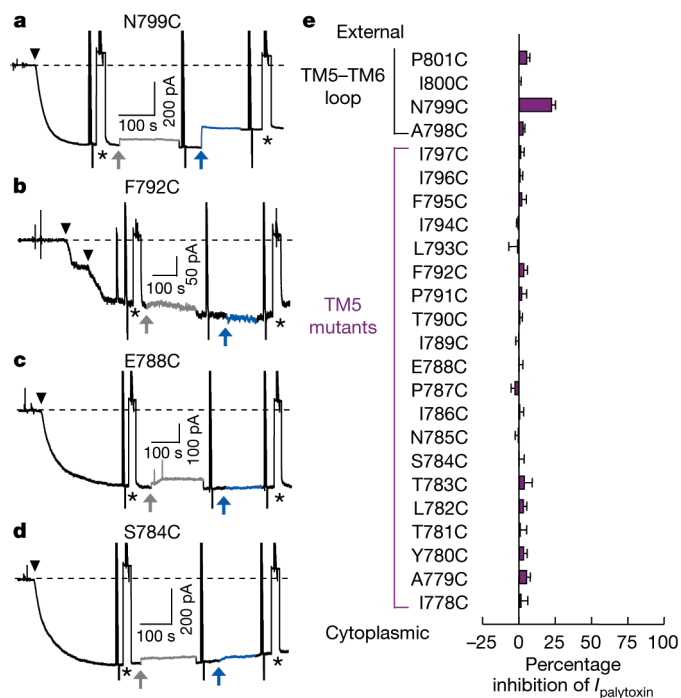


Figure 2 | Effects of MTSET^+ on current through palytoxin-bound Na^+, K^+ pump-channels with cysteines in TM5 or the TM5–TM6 loop. a–d, Current at -50 mV in outside-out patches exposed to symmetrical Na^+ concentrations. Application of 50 nM palytoxin (black arrowheads) generated inward (negative) current, $I_{\text{palytoxin}}$ (dashed line marks zero total membrane current). Temporary substitution (asterisk) of less permeant tetramethylammonium (TMA^+) for external Na^+ monitored patch integrity. Application of 10 mM dithiothreitol (grey arrows, grey traces) caused a small, reversible, poorly understood current decrease. Then 1 mM MTSET^+ (blue arrows, blue traces) was applied until the current became steady. e, Summary of mean ($\pm \text{s.e.m.}$; n , 3–6 patches) percentage inhibition of $I_{\text{palytoxin}}$ by 1 mM MTSET^+ at -50 mV for each single-cysteine mutant.

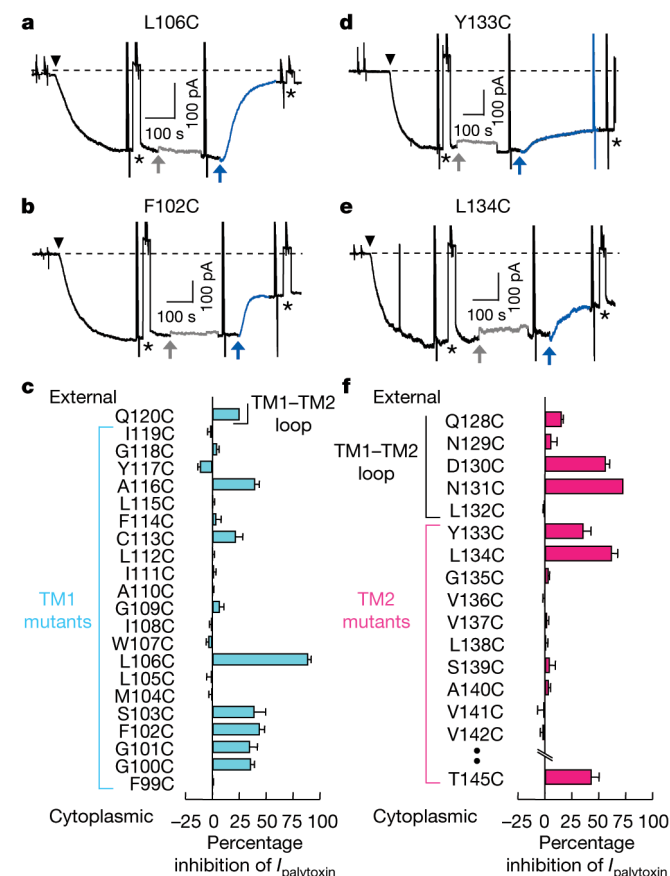


Figure 3 | Effects of MTSET^+ on current through palytoxin-bound Na^+, K^+ pump-channels with cysteines in TM1, TM2 or the TM1–TM2 loop. a, b, d, e, Representative current recordings in outside-out patches under the same conditions, and with the same applications of palytoxin, TMA^+ , dithiothreitol and MTSET^+ , as in Fig. 2. c, f, Summary of percentage inhibition of $I_{\text{palytoxin}}$ by 1 mM MTSET^+ at -50 mV for each single-cysteine mutant, given as mean ($\pm \text{s.e.m.}$; n , 3–11 patches, except for Q120C ($n = 2$) and N131C ($n = 1$), both previously shown¹⁸ to be MTSET^+ accessible). C113C indicates data from wild-type, ouabain-sensitive *Xenopus* Na^+, K^+ pumps tested (in the absence of ouabain) in patches from non-injected control oocytes.

(MTSES⁻), despite reaction with similarly sized, but positively charged, MTSET⁺ (ref. 5). MTSET⁺ reaction with cysteines substituted for deeper TM1 residues L106, S103, F102, G101 and G100 (Fig. 4c) decreased current by ~40–90% (Figs 3a–c, 4c). The smaller current decrease, of 20–30%, on reaction with the comparably sized neutral reagent 2-aminocarbonyl-ethyl-methanethiosulphonate (MTSACE; Fig. 4c and Supplementary Fig. 4) is consistent with simple steric interference with Na⁺ current flow by the ~6 Å × 8 Å adduct. Negatively charged MTSES⁻, however, failed to react, neither altering pump–channel current nor preventing its subsequent decrease by MTSET⁺ (Fig. 4c; Supplementary Fig. 4a, c); deep TM2 position T145 (see Fig. 3f) behaved comparably. By contrast, MTSES⁻ increased current in pump–channels with cysteines at the more superficial TM1 position A116 and TM1–TM2 loop residues Q128 and D130 (as previously shown¹⁸ for Q120 and N131), the negative adduct electrostatically elevating the local concentration of current-carrying Na⁺ ions^{5,18}. These results show that MTS reagents had to pass the cation-selectivity filter formed by E336 and D813 to reach every deeper reactive cysteine.

Our findings are all broadly consistent with corresponding locations of target residues in the Na⁺,K⁺ pump model based on the SERCA E2·BeF₃⁻ structure (Figs 1b, 4a, b; Supplementary Figs 1, 2), supporting its overall applicability. This is despite both the mere 26% amino acid identity between SERCA and Na⁺,K⁺-ATPase in the TM1–TM6 region scanned here, and the fact that the cytoplasmic-side pathway is tightly^{12,13} closed in SERCA E2·BeF₃⁻ (and also, apparently, in Na⁺,K⁺-ATPase E2·BeF₃⁻; Supplementary Fig. 5), whereas it can demonstrably open in the Na⁺,K⁺ pump–channel. That open cytoplasmic access pathway runs between TM1, TM2 and

TM4, beyond the TM1 kink (at G101) seen in E2 structures^{10,11} (Fig. 4b; Supplementary Figs 1, 2). MTSET⁺-accessible TM1 positions G101, F102, and L106 (Figs 3, 4) correspond to residues (rat Na⁺,K⁺-ATPase α₁ G94, F95 and L99) important in Na⁺ and K⁺ binding and occlusion in E2 conformations^{25,26}, with L99 (here L106) in particular²⁶ cooperating with E329 (here E336) to lock exit or entry at site II. The equivalent SERCA TM1 region appears to gate cytoplasmic access for Ca²⁺ ions^{8,10,11}.

We found no sign of reaction with 1 mM MTSET⁺ at any of 20 contiguous TM5 positions (Fig. 2), even though the E788-equivalent TM5 residue appears accessible from the extra-cytoplasmic side in the open¹² SERCA E2·BeF₃⁻ structure (Supplementary Fig. 6), and residues equivalent to S784, N785 and E788 help coordinate Rb⁺ at site I in the occluded E2·MgF₄²⁻ Na⁺,K⁺ pump when both gates are shut²¹. Given that MTSET⁺ reaction at many nearby positions altered pump–channel current (Fig. 4), it is unlikely that TM5 sites reacted without modifying current.

Although we cannot rule out the possibility that distortion of the Na⁺,K⁺-ATPase ion pathway by palytoxin made TM5 residues inaccessible, this seems unlikely for several reasons. First, palytoxin action can be readily reversed, and repeated, on the same population of Na⁺,K⁺ pumps³. Second, the gates to the ion pathway through palytoxin-bound pump–channels still respond to the Na⁺,K⁺ pumps' physiological ligands³. Third, positions as deep as the pathway narrowing are accessible to MTSET⁺ without palytoxin (Supplementary Fig. 3). Fourth, blockers of access channels to ion-binding sites in unmodified Na⁺,K⁺ pumps similarly impede cation movement in palytoxin-bound pumps²³. Fifth, MTSET⁺-accessible positions in palytoxin-bound pump–channels map reasonably onto unmodified pump structures (Fig. 4a, b and Supplementary Figs 1, 2; compare with refs. 5, 16, 17) and include sites expected to interact with transported ions^{19–21,25,26}.

We conclude that unfavourable geometry precluded reactivity of MTSET⁺ with TM5 positions at site I because they do not lie on the principal ion pathway. This is consistent with the side-chain charge of site-I residues E778 and D817 having little apparent influence on cation selectivity of Na⁺,K⁺ pump–channels⁵. It is also consistent with the very slow reaction of E788C with a smaller reagent, namely 1-trimethylammonium-methyl-methanethiosulphonate (MTSMT⁺; Supplementary Fig. 7); similarly small 2-aminoethyl-methanethiosulphonate (MTSEA⁺; compare with ref. 15) is unreliable as it is membrane permeant and slowly reacts with Na⁺,K⁺ pumps lacking engineered cysteines (Supplementary Fig. 7).

That ~6 Å wide, ~12 Å long MTSET⁺, MTSES⁻ and MTSACE pass through palytoxin-bound Na⁺,K⁺ pump–channels corroborates the findings that these channels conduct *N*-methyl-D-glucamine ions (diameter ≥7 Å) only ~50 times more slowly than Na⁺ ions²², and that their measured²⁷ Na⁺ flux ratio exponent²⁸ is ~1.0, implying little interaction between Na⁺ ions in a queue along the principal pathway that passes through site II. Occupancy by a second Na⁺ ion of site I, off the main pathway but linked to it by a connection narrow enough to preclude reactivity with MTSET⁺, could account for suggested average pump–channel occupancy by two Na⁺ ions²³. In SERCA, lock-in of a Ca²⁺ ion in site I by binding of the second Ca²⁺ ion²⁹ in site II, and sequential release of the two Ca²⁺ ions³⁰, are similarly consistent with transported ions negotiating a single common pathway from the cytoplasm to the ion-binding sites in E1 states, and from those sites to the reticulum lumen during release in the E2P state. The present snapshot of an ion pathway right through the Na⁺,K⁺ pump affords a structural basis for understanding cation translocation in P-type pumps.

METHODS SUMMARY

Ouabain- and MTS-insensitive *Xenopus* Na⁺,K⁺ pumps. *Xenopus* Na⁺,K⁺ pumps were made insensitive to ouabain and extracellular MTS reagents by the mutation C113Y (ref. 24) in *Xenopus* Na⁺,K⁺-ATPase α₁ subunits as described previously¹⁸. Single cysteines were introduced into C113Y Na⁺,K⁺-ATPase α₁ by

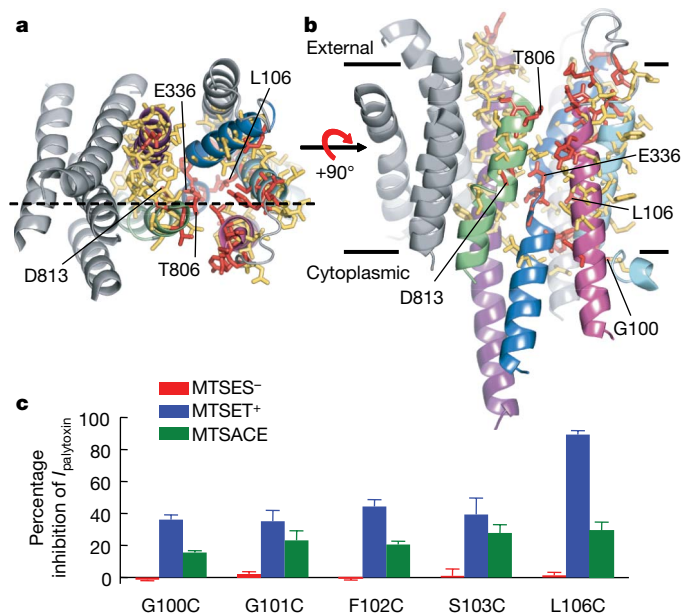


Figure 4 | Structural model and characteristics of ion pathway through the palytoxin-bound Na⁺,K⁺-ATPase. Results (including reactive and non-responsive positions from ref. 5) mapped onto a homology model of the Na⁺,K⁺-ATPase transmembrane domain (helices coloured as in Fig. 1) based on the SERCA E2·BeF₃⁻ structure¹², viewed from the extracellular surface (a) or from the membrane plane (b). Dashed line in a indicates plane of cut in Supplementary Fig. 2a. Red sticks mark reactive positions (*I*_{palytoxin} altered by >10% by MTSET⁺) and yellow sticks mark non-responsive positions. Reaction rate constants for MTSET⁺ decreased from ≥10⁴ M⁻¹s⁻¹ for superficial positions to ≥10 M⁻¹s⁻¹ for deep positions (Supplementary Fig. 8). c, Accessibility of cysteines beyond the cation-selectivity filter depends on the charge of the MTS reagent; summary of mean percentage inhibition (± s.e.m.; *n*_s 3–8 patches) of *I*_{palytoxin} at ~50 mV by ~2.5-min applications (all 1 mM) of MTSES⁻ (red bars), MTSET⁺ (blue bars) or MTSACE (green bars).

PCR. Complementary DNA in a pSD5 vector was transcribed *in vitro*. *Xenopus* oocytes were injected with a 50-nl mixture of 5 ng of *Xenopus* β_3 and 15 ng of mutated *Xenopus* α_1 complementary RNAs, and incubated at 18 °C for 1–3 days.

Current recordings and analysis. Currents were recorded in outside-out excised patches at 22–24 °C as described previously^{5,18}. The internal (pipette) solution contained 125 mM NaOH, 100 mM sulphamic acid, 20 mM HCl, 10 mM HEPES, 1 mM EGTA, 1 mM MgCl₂ and 5 mM MgATP (pH 7.4). The external solution contained 125 mM NaOH or TMA-OH, 125 mM sulphamic acid, 10 mM HEPES, 5 mM BaCl₂, 0.5 mM CaCl₂, 1 mM MgCl₂ (pH 7.6) and 100 μ M ouabain. Palytoxin (Wako) was added (from 100 μ M aqueous stock solution) at 50 nM, with 0.001% bovine serum albumin and 1 mM Na-borate. MTS reagents (Toronto Research Chemicals) were added from ice-cold (~0 °C) 100 mM aqueous stock solutions immediately before use, and were refreshed at 1.5-min intervals to maintain reactivity during prolonged (≥ 2 min) applications⁴. Alteration of palytoxin-induced current by MTS reagents was calculated as follows: percentage inhibition of $I_{\text{palytoxin}}$ equals $100 \times (1 - I_{\text{after}}/I_{\text{before}})$. Here I_{after} represents steady palytoxin-induced current at –50 mV after MTS reagent application, and I_{before} represents the same current just before MTS reagent application. Data are given as mean \pm s.e.m.

Model building. The *Xenopus* Na⁺,K⁺-ATPase α_1 subunit homology model was built from the Ca²⁺-ATPase E2·BeF₃[–] structure (ref. 12; Protein Data Bank code 3B9B) using SWISS-MODEL (<http://swissmodel.expasy.org>) as described previously⁵. Structural figures were prepared with PyMOL version 0.97 (<http://www.pymol.org>).

Received 14 April; accepted 20 August 2008.

Published online 8 October 2008.

- Läuger, P. *Electrogenic Ion Pumps* (Sinauer, 1991).
- Scheiner-Bobis, G., Meyer zu Heringdorf, D., Christ, M. & Habermann, E. Palytoxin induces K⁺ efflux from yeast cells expressing the mammalian sodium pump. *Mol. Pharmacol.* **45**, 1132–1136 (1994).
- Artigas, P. & Gadsby, D. C. Na⁺/K⁺-pump ligands modulate gating of palytoxin-induced ion channels. *Proc. Natl Acad. Sci. USA* **100**, 501–505 (2003).
- Karlin, A. & Akabas, M. H. Substituted-cysteine accessibility method. *Methods Enzymol.* **293**, 123–145 (1998).
- Reyes, N. & Gadsby, D. C. Ion permeation through the Na⁺,K⁺-ATPase. *Nature* **443**, 470–474 (2006).
- Toyoshima, C., Nomura, H. & Tsuda, T. Lumenal gating mechanism revealed in calcium pump crystal structures with phosphate analogues. *Nature* **432**, 361–368 (2004).
- Toyoshima, C., Nakasako, M., Nomura, H. & Ogawa, H. Crystal structure of the calcium pump of sarcoplasmic reticulum at 2.6 Å resolution. *Nature* **405**, 647–655 (2000).
- Toyoshima, C. & Nomura, H. Structural changes in the calcium pump accompanying the dissociation of calcium. *Nature* **418**, 605–611 (2002).
- Olesen, C. *et al.* Dephosphorylation of the calcium pump coupled to counterion occlusion. *Science* **306**, 2251–2255 (2004).
- Sørensen, T. L. M., Møller, J. V. & Nissen, P. Phosphoryl transfer and calcium ion occlusion in the calcium pump. *Science* **304**, 1672–1675 (2004).
- Toyoshima, C. & Mizutani, T. Crystal structure of the calcium pump with a bound ATP analogue. *Nature* **430**, 529–535 (2004).
- Olesen, C. *et al.* The structural basis of calcium transport by the calcium pump. *Nature* **450**, 1036–1042 (2007).
- Toyoshima, C. *et al.* How processing of aspartylphosphate is coupled to luminal gating of the ion pathway in the calcium pump. *Proc. Natl Acad. Sci. USA* **104**, 19831–19836 (2007).
- Sakmann, B. & Neher, E. *Single-Channel Recording* (Plenum, 1995).
- Guennoun, S. & Horisberger, J. D. Structure of the 5th transmembrane segment of the Na,K-ATPase α subunit: a cysteine-scanning mutagenesis study. *FEBS Lett.* **482**, 144–148 (2000).
- Guennoun, S. & Horisberger, J. D. Cysteine-scanning mutagenesis study of the sixth transmembrane segment of the Na,K-ATPase α subunit. *FEBS Lett.* **513**, 277–281 (2002).
- Horisberger, J. D., Kharoubi-Hess, S., Guennoun, S. & Michielin, O. The fourth transmembrane segment of the Na,K-ATPase α subunit: a systematic mutagenesis study. *J. Biol. Chem.* **279**, 29542–29550 (2004).
- Artigas, P. & Gadsby, D. C. Ouabain affinity determining residues lie close to the Na/K pump ion pathway. *Proc. Natl Acad. Sci. USA* **103**, 12613–12618 (2006).
- Nielsen, J. M., Pedersen, P. A., Karlsh, S. J. & Jorgensen, P. L. Importance of intramembrane carboxylic acids for occlusion of K⁺ ions at equilibrium in renal Na,K-ATPase. *Biochemistry* **37**, 1961–1968 (1998).
- Ogawa, H. & Toyoshima, C. Homology modeling of the cation binding sites of Na⁺K⁺-ATPase. *Proc. Natl Acad. Sci. USA* **99**, 15977–15982 (2002).
- Morth, J. P. *et al.* Crystal structure of the sodium-potassium pump. *Nature* **450**, 1043–1049 (2007).
- Artigas, P. & Gadsby, D. C. Large diameter of palytoxin-induced Na/K pump channels and modulation of palytoxin interaction by Na/K pump ligands. *J. Gen. Physiol.* **123**, 357–376 (2004).
- Harmel, N. & Apell, H. J. Palytoxin-induced effects on partial reactions of the Na,K-ATPase. *J. Gen. Physiol.* **128**, 103–118 (2006).
- Canessa, C. M., Horisberger, J. D., Louvard, D. & Rossier, B. C. Mutation of a cysteine in the first transmembrane segment of Na,K-ATPase α subunit confers ouabain resistance. *EMBO J.* **11**, 1681–1687 (1992).
- Einholm, A. P., Toustrup-Jensen, M., Andersen, J. P. & Vilsen, B. Mutation of Gly-94 in transmembrane segment M1 of Na⁺,K⁺-ATPase interferes with Na⁺ and K⁺ binding in E₂P conformation. *Proc. Natl Acad. Sci. USA* **102**, 11254–11259 (2005).
- Einholm, A. P., Andersen, J. P. & Vilsen, B. Importance of Leu⁹⁹ in transmembrane segment M1 of the Na⁺,K⁺-ATPase in the binding and occlusion of K⁺. *J. Biol. Chem.* **282**, 23854–23866 (2007).
- Rakowski, R. F. *et al.* Sodium flux ratio in Na/K pump-channels opened by palytoxin. *J. Gen. Physiol.* **130**, 41–54 (2007).
- Hodgkin, A. L. & Keynes, R. D. The potassium permeability of a giant nerve fibre. *J. Physiol. (Lond.)* **128**, 61–88 (1955).
- Zhang, Z., Lewis, D., Strock, C. & Inesi, G. Detailed characterization of the cooperative mechanism of Ca²⁺ binding and catalytic activation in the Ca²⁺ transport (SERCA) ATPase. *Biochemistry* **39**, 8758–8767 (2000).
- Inesi, G. Sequential mechanism of calcium binding and translocation in sarcoplasmic reticulum adenosine triphosphatase. *J. Biol. Chem.* **262**, 16338–16342 (1987).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank N. Fataliev for help with molecular biology, the late R. F. Rakowski for cDNAs encoding *Xenopus* α_1 and β_3 Na⁺,K⁺-ATPase subunits, and P. Nissen, B. Vilsen and J. V. Møller for providing atomic coordinates before their publication. The work was supported by a grant from the NIH (to D.C.G.) and a fellowship from the Vicente Trust (to P.A.); N.R. is presently a Jane Coffin Fund Fellow. We dedicate this paper to the memory of our colleague R. F. Rakowski.

Author Information Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to D.C.G. (gadsby@rockefeller.edu).

naturejobs

**THE CAREERS
MAGAZINE FOR
SCIENTISTS**

What's the fairest way to allot precious government funds to promising biomedical scientists? In an initiative developed by its outgoing director Elias Zerhouni, the US National Institutes of Health (NIH) has decided that ensuring a minimum level of support for new investigators should be part of an equitable formula — one that benefits both the investigator and the scientific enterprise.

In late October, the NIH made official a policy that has been under review for the past two years. It aims to ensure that grant applicants who have never received NIH funding should have success rates comparable to those who have received NIH funds at least once. The agency will achieve this by routinely funnelling the applications of new investigators into a separate pool initially so that they are not competing with more experienced NIH grantees. For the fiscal year 2009, this translates into at least 1,650 awards to new investigators across the NIH's centres.

The move is meant to reduce reviewer bias against first-time applicants, a documented trend that, unsurprisingly, sees existing and previous NIH grantees — who have demonstrated success and can provide more data to buttress their proposals — typically faring better.

But Zerhouni's move was not inevitable, obvious or risk-free. With this additional attention towards first-timers, grantees with more experience — or, indeed, with just a single NIH grant — have slightly diminished chances. Is Zerhouni unfairly gaming the system? And will the quality of the science suffer as a result? It is possible, in principle. But with the average age of the new NIH grantee on the rise — up to 42 from 37 in 1980 — Zerhouni has recognized that exceptional young biomedical scientists must have a reasonable chance to succeed, and that new ideas and approaches are essential for a healthy research enterprise. At the very least, the policy sends a signal to eager but discouraged young scientists — part of its goal, says the NIH's deputy director of extramural research, Sally Rockey. "It's a way for them to understand that the NIH is possible."

Gene Russo is editor of *Naturejobs*.

CONTACTS

Editor: Gene Russo

European Head Office, London
The Macmillan Building,
4 Crinan Street, London N1 9XW, UK
Tel: +44 (0) 20 7843 4961
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

European Sales Manager:
Andy Douglas (4975)
e-mail: a.douglas@nature.com
Assistant European Manager:
Nils Moeller (4953)

Natureevents:
Ghizlaine Ababou (+44 (0) 20 7014 4015)
e-mail: g.ababou@nature.com

Southwest UK/RoW:
Alexander Ranken (4944)

Northeast UK/Ireland:
Matthew Ward (+44 (0) 20 7014 4059)

France/Switzerland/Belgium:
Muriel Lestringuez (4994)
Scandinavia/Spain/Portugal/Italy:
Evelina Rubio-Hakansson (4973)
North Germany/The Netherlands/Eastern Europe: Kirstin Vincze (4970)
South Germany/Austria:
Hildi Rowland (+44 (0) 20 7014 4084)

Advertising Production Manager:
Stephen Russell
To send materials use London address above.
Tel: +44 (0) 20 7843 4816
Fax: +44 (0) 20 7843 4996
e-mail: naturejobs@nature.com

Naturejobs web development: Tom Hancock
Naturejobs online production: Dennis Chu

US Head Office, New York
75 Varick Street, 9th Floor,
New York, NY 10013-1917
Tel: +1 800 989 7718

Fax: +1 800 989 7103
e-mail: naturejobs@natureny.com

US Sales Manager: Peter Bless

India
Vikas Chawla (+91 1242881057)
e-mail: v.chawla@nature.com

Japan Head Office, Tokyo
Chiyoda Building, 2-37 Ichigayatamachi,
Shinjuku-ku, Tokyo 162-0843
Tel: +81 3 3267 8751
Fax: +81 3 3267 8746

Asia-Pacific Sales Manager:
Ayako Watanabe (+81 3 3267 8765)
e-mail: a.watanabe@natureasia.com
Business Development Manager, Greater China/Singapore:
Gloria To (+852 2811 7191)
e-mail: g.to@natureasia.com

MOVERS

William Brody, president, Salk Institute for Biological Studies, La Jolla, California



1996–2008: President, Johns Hopkins University, Baltimore, Maryland
1994–96: Provost, Academic Health Center, University of Minnesota, Minneapolis
1987–94: Professor of radiology, biomedical engineering and electrical and computer engineering, Johns Hopkins University

When William Brody decided to step down as president of Johns Hopkins University in Baltimore, Maryland, earlier this year, the 64-year-old was looking forward to some downtime. But just as he was about to publicly announce his retirement, the Salk Institute came calling. "I asked if they were tapping my phone," says Brody, who will start as Salk's president next March.

Brody had assumed Salk was looking for an active molecular biologist. But Salk didn't want someone brought from the bench to manage fellow scientists. The small institute, often described as talent-rich but resource-poor, wanted a big name that could bolster its small endowment. Salk also wanted some stability; in the past two decades most of its presidents have stayed for only a year or two.

Brody fits the bill. Trained as a PhD-MD, he was Hopkins's president for 12 years and says he has made a five-year commitment to Salk. He will bring the connections that helped him complete a US\$3.2-billion fundraising campaign at Hopkins. But the Salk Institute is no Hopkins, which spends \$1.5 billion a year on research and development — the most in the United States — and has some 44,000 researchers. Salk has 870 researchers, and acquires most of its \$113-million budget from federal, state and philanthropic grants. "Clearly we have to build up the resource base — which is going to be much harder in this economic environment," Brody says.

It will be a sort of homecoming for Brody, who grew up in Stockton, California. After pursuing electrical engineering at the Massachusetts Institute of Technology, he earned his PhD-MD from Stanford University in California. He put those skills to use in the field of radiology, developed several patents, and co-founded three medical-device companies. At Hopkins, he was one of the best-paid university presidents, making \$1.9 million in 2005–06, according to a 2007 *Chronicle of Higher Education* report. At Salk, he says, "it'll be a whole lot less".

Salk will get an excellent president and fundraiser for a bargain price, says Ruth Faden, director of the Berman Institute of Bioethics at Johns Hopkins. Faden recalls a meeting between Brody — a renowned pianist — and benefactor Phoebe Berman, when she was contemplating a multimillion-dollar gift to found the institute. "I want to make sure you're everything you claim. Can you play?" Faden recalls Berman saying. "Bill didn't bat an eyelash, and sat down and played," says Faden. "She said, 'Okay, you got it.'" ■
Eric Hand

NETWORKS & SUPPORT

Building up to an HIV vaccine

Vaccine development is a notoriously challenging career path, given the propensity for negative results. For every 100 potential immunogens made, very few will work. Young scientists often choose research areas more likely to offer career advancement. Hoping in part to address these shortcomings, the International AIDS Vaccine Initiative (IAVI), a global non-profit organization, is investing \$30 million to create a new HIV Neutralizing Antibody Center at the Scripps Research Institute in La Jolla, California.

"One of the major challenges in the HIV world is bringing young investigators into the field and maintaining their career path," says Wayne Koff, IAVI's senior vice-president of research and development. "We want to train the next generation of HIV vaccine discovery scientists." The organization is seeking young scientists willing to commit to a multi-year programme. It intends to combine expertise — at the graduate student, postdoc and scientist level — in immunology, molecular biology, protein chemistry, molecular virology, computational biology and drug discovery.

"The recent failure of the Merck HIV vaccine highlighted the need

to think about antibodies as well the T-cell component in HIV vaccine development," says Scripps immunologist Dennis Burton, who is scientific director of the HIV Neutralizing Antibody Consortium (NAC), a collaboration between the IAVI and leading AIDS laboratories in the United States and Europe.

The centre plans to promote a multidisciplinary approach and dedicate time to mentoring. Graduate students will be afforded the chance to develop soft skills, such as grant-writing and giving presentations at international meetings, at a much earlier stage than is usual in academic settings. Young scientists will have instant connections to the top HIV labs around the world through the NAC.

Initially the centre will have 30 people, says Koff, including some senior international scientists within the consortium. The IAVI hopes to have the centre under way by the beginning of 2009. "This is the highest vaccine-discovery priority at the IAVI," says Koff. He says Scripps would like to identify an immunogen consisting of broadly neutralizing antibodies against HIV and be on its way to clinical development of that immunogen within five years. ■

Virginia Gewin

POSTDOC JOURNAL

Imposter syndrome

Last week, while giving a seminar, I suddenly felt as if the sentences coming out of my mouth weren't mine. And the brain controlling that mouth didn't seem to be mine either. This was not the first time I had suffered 'imposter syndrome'.

Although not officially recognized as a psychological disorder, it might have been if doctors had used me as a case study a few years back. Whether I had success in sub-cloning a long DNA fragment or fashioning a unique experimental design, I would usually attribute such accomplishments to luck.

I was relieved to discover that imposter syndrome is widespread in academia: sometimes successful people are unable to accept that their accomplishments are deserved. In extreme cases, I've heard it can lead people to avoid challenges altogether, for fear of failure. Fortunately, I have never suffered to that extent.

In fact, what I have come to realize is that the syndrome is a mechanism to keep my ego in check. This can be a good thing. An ego can be a dangerous monster, as I found out as an undergraduate student when I alienated a few lab mates by gloating that my fruitfly genetics project was far superior to theirs. Ironically, not long after, I scored a paltry 65 on my fruitfly genetics exam. Perhaps a few recent successes explain the resurgence of the imposter in my seminar. Maybe it's time my ego was tamed once again. ■

Zachary Lippman is a postdoctoral fellow at the Hebrew University of Jerusalem's faculty of agriculture in Israel.

Birthday surprises

Now you are 16.

Erika Cule

Strains of *Happy Birthday* greeted me from the kitchen.

"... Happy Birthday dear Elizabeth, Happy Birthday to you!"

I grinned and embraced my father. "Many happy returns, love," he said. He ruffled my hair affectionately. "Sweet sixteen, quite something."

Grandma beckoned from her armchair. "Come and give your Nan a kiss."

I rushed to hug her. She tried to hide a wince as I brushed against her scars. "Your future should be in your inbox now," she said, with a nod towards my laptop. "I asked for it to be delivered this morning."

Thoughts of my father's gift forgotten, I snatched up my machine and made for our neighbourhood coffee shop. I needed some privacy while I studied what life had in store. Installed on a sofa with a coffee, I logged on. Among the dozens of birthday greetings was the message from sequencing@545.com.

Subject: Your Complete Genome Sequence
Birthday Greetings, Elizabeth.

Thank you for ordering Your Complete Genome Sequence High-Quality Version from 545.com. Our Annotation Service supplies a detailed genotype analysis. We analysed more than 6 million single nucleotide polymorphisms, alongside mini- and micro-satellites, deletions, inversions and copy number variations. We report known phenotypic associations, together with significance levels. All associations reference peer-reviewed publications and are cross-referenced to GenBank and Online Mendelian Inheritance in Man where relevant. Your annotated sequence has been uploaded to the Worldwide Depository of Genomes. It can be identified by your unique GenomeID, which you are free to distribute to potential employers, insurers or acquaintances, at your discretion.

I caught my breath and took a gulp of coffee. I had been waiting for this since 545 launched its service two years ago. My grandmother, a geneticist by training, was as excited as I was about reading my future.

I opened the Annotation file.

Personality and Aptitudes

I sought out the details I needed.

Estimated IQ: 138.

Critical and Analytical

Ability: Ninety-eighth percentile

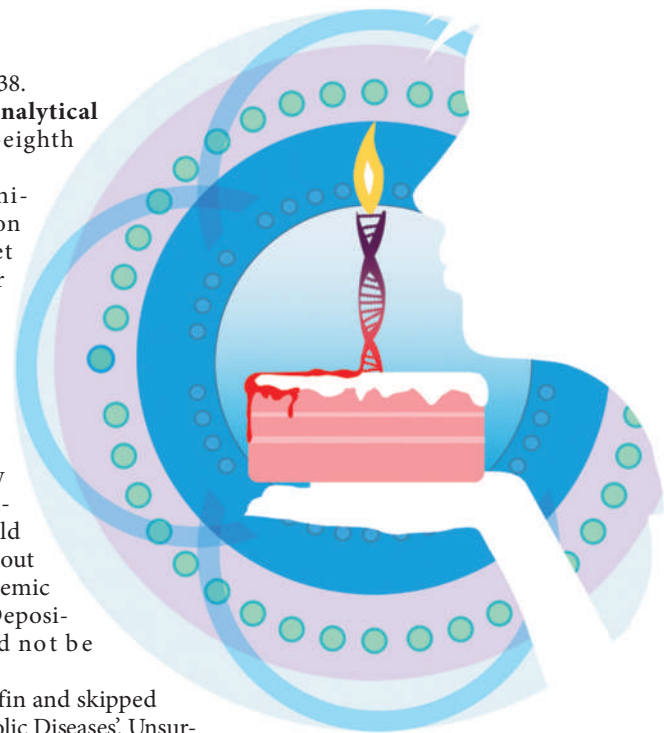
Next term's university application form did not yet have a space for Genome ID. However, it would not hurt to drop it into my Personal Statement to illustrate my curiosity about genetics. I knew that Imperial College London would have no qualms about checking my academic potential in the Depository. They would not be disappointed.

I ordered a muffin and skipped through to 'Metabolic Diseases'. Unsurprisingly, a set of seven SNPs inferred a haplotype strongly associated with type 2 diabetes. With three of my grandparents, together with the obese half of the population, now falling into that category, the emerging consensus among endocrinologists was that the condition represented the wild-type genotype of the twenty-first century. I pinched my waistline. The muffin lost its appeal.

Appearance: Homozygosity for the BN4 allele, associated with poorly regulated cartilage formation in the nasal region.

Advertisements for rhinoplasty scrolled across the bottom of the screen. Idly, I clicked on one, wondering whether surgery would have time to heal before university started. If you cannot change nature, nurture would be on hand to help out.

'Oncogenes and Late Onset Disorders' did not surprise me in reporting homozygosity for a particular variant in *BRCA2*. My grandmother had been one of the last patients to be treated surgically, a decade ago. The latest treatment was injected engineered leukocytes that homed in on metabolically hyperactive tumours and restricted angiogenesis, claiming to 'starve your tumours before they consume you'. This was the technology that inspired me to apply for the 'Nanobots and Cellular Machines' course at Imperial, which was refining the method and developing more applications. I drifted



into a daydream where I was head of a spin-out from Imperial Innovations, jetting to conferences with a briefcase of nanobots.

My cellphone buzzed — Dad, reminding me about our lunch date. I had been absorbed in my future for the entire morning. It did not look too bad. No Huntington's, no Alzheimer's, no severe immune defects.

No nasty genetic surprises.

I handed a five-euro note to the barista. "Keep the change," I smiled.

"What's the good news?" He raised an eyebrow flirtatiously.

"I'm not going to get struck down by an inheritable disorder anytime soon."

"Fantastic," he mused, "although everyone's got to die sometime."

The receipt he gave me had something scrawled at the bottom. I squinted as I stepped out of the coffee shop and into the sunlight. It was a telephone number. Maybe my poorly regulated nasal cartilage was not so bad.

Neither he, I nor my genome noticed the oncoming bus.

Erika Cule is a recent graduate in biochemistry from Imperial College, London. This article was the winning adult entry in the Daily Telegraph/Bayer 2008 Young Science Writers competition (www.science-writer.co.uk) and is reproduced courtesy of *The Daily Telegraph*.

JACEY